ON THE TREATMENT OF TEST ANXIETY: ITS NATURE, MEASUREMENT, EVALUATION, AND SELF-EFFICACY

FRANCESCO URSINO

A thesis
submitted in partial fulfillment
of the requirements for the Degree of
MASTER OF SCIENCE IN PSYCHOLOGY

University of Canterbury

February 1986

To my mother

ACKNOWLEDGEMENTS

I am indebted to all the subjects who took part in this study and whose perseverance during treatment and patience in completing the large number of questionnaires has made this thesis possible.

I would like to thank my supervisors Dr. Nirbhay Singh and Dr. Roger Katz for their help.

I am greatly indebted to Dr. Kenneth Holroyd of the University of Miami (USA) for a copy of his treatment manual and to Dr. Garth Ritchie for his help and suggestions relating to the data analysis.

I am very greatful to Dr. Hugh Lauder for his help and encouragement during the early stages of this study.

The large number of measures employed in this study has needed the cooperation of a number of people. I would like to thank the kind help of: Prof. Albert Bandura of Stanford University (USA), Drs. Suzanne Bennett and Gina Harris of the University of Florida (USA), Dr. Robert Osterhouse of Prince George's Community College (Maryland, USA), Dr. Samuel Osipow of the Ohio State University (USA), Mr. Bruce Jamieson, Ms. Elizabeth Gutteridge, and the Education Department test library staff.

I would like to express my appreciation of the service offered by the Library and particularly the friendly and efficient service of the circulation desk staff.

I am very grateful to Maire Kipa for typing the original manuscript with such delightful dedication and to Andrew Ivory

for proof-reading so well.

I would like to express my sincere appreciation to the tremenduous help given by Ijan Beveridge even in a time when he himself was very busy. I am equally appreciative of the assistance given by Dr. Steve Hudson and Mr. Neville Blampied.

In sharp contrast to the writing of this thesis, the editing and printing of it in time for the final deadline has been very much a collective effort, for I am very grateful to Tessa Copland, Kevin Moesbergen, Janine Watson, and Alister Swale.

ABSTRACT

Test anxiety is a very complex phenomenon. It has been argued that the reason why comparative test anxiety treatment studies have usually reported negative or inconsistent results is due to (1) poor evaluation procedures and (2) using test anxiety as a convenient arena for testing psychotherapy techniques rather than attempting to devise a treatment that would effectively decrease test anxiety and increase academic performance. Based upon extensive research evidence an appropriate framework was developed within which to test the effectiveness of a treatment package (cognitive-attentional training plus study and exams skills) for test anxiety. The results showed this treatment package to be very effective in reducing exam anxiety. A vigorous effort was made to understand the course of test anxiety treatment. General anxiety, achievement motivation, verbal ability, numerical ability, and ethnocentrism affected the course of treatment at decreasing test anxiety, on one or more of the measures, in the short or in the long term or both. State test anxiety and Studying Habits (SH) affected the course of treatment in decreasing general anxiety. Subjects in the broad middle range of the numerical ability distribution had their cognitive-intellectual performance most affected by test anxiety and by changing levels of test anxiety. While greater test anxiety decreases were associated with greater increases in the more valid academic performance measure, they were associated with greater dercreases in performance on another academic measure as well as three ability meaures. A distinction was made within self-efficacy theorizing.

Self-efficacy in managing one's anxiety (SEMA) and self-efficacy in succeeding (SESC). SEMA is tantamount to the converse of anxiety. Two new measures were developed and validated (TASEQ and P-TASEQ) to enable the empirical testing of such distinction. Success of therapy paralleled subjects' levels of self-efficacy feelings.

CONTENTS

INTRODUCTION

"We live in a test-conscious, test giving culture in which the lives of people are in part determined by their test performance". This observation comes from the writing of Seymour Sarason (1959, p. 26) a prominent figure in this area of research, who with George Mandler (Mandler & Sarason, 1952; S. Sarason & Mandler, 1952) provided the impetus for theory and research in the area of test anxiety. Of course, if we apply the above quote to school or university life we soon realise that it becomes an understatement: the lives of high school pupils and university students are strongly determined by their test performance.

A brief examination of the Japanese educational system shows that indeed in this culture performance on one exam (the university entrance exam) has far reaching consequences on the lives of Japanese people.

Whereas in western cultures the educational evaluation of an individual is spread over many years, "a Japanese person's entire career depends upon his success in a simple event, the entrance examination to university" (Iga, 1979, p. 22). The severity of the education system is intensified by the way educational institutions are ranked: by the number of graduates admitted to "better" universities. Newspapers and magazines publish such ranking. Moreover, the ranking corresponds to the expected career success of graduates because more prestigious employers hire graduates only from highly ranked institutions.

To make this system even more rigid, the more highly

ranked universities are the governmental ones whose tuitions are only about one tenth of those of private universities (Iga, 1979). Even the governmental universities are ranked and those in the higher echelons get more and better applicants than those below them, this fact adds an even stronger pyramidal flavour to this extremely rigid structure.

So much importance is placed on the university entrance exam that it creates severe pressures on the life of Japanese students and their families. The stresses produced by Japan's examination system, often referred to as the "Examination Hell" (shiken jigoku), has been blamed as the main cause of suicide of Japanese youth which has been one of the highest in the world (Iga, 1979).

It appears that the greater the emphasis on one exam the more deleterious the effects of the stress it creates on people. To this point Brown (1938a) found that "seniors at schools employing one final comprehensive exam" were the most "excited" group of his subjects as measured by a questionnaire which was essentially the first test anxiety scale reported in the literature. On the other hand, "seniors where exams were on final quarterly basis" were the least "excited" group of his subjects. More recent evidence (e.g., Morris & Fulmer, 1976; I. Sarason, 1972; S. Sarason, Davidson, Lighthall, Waite, & Ruebush, 1960) also suggests that the more important the test the greater the test anxiety it evokes.

Luckliy in western educational systems the emphasis placed on one exam is never as great and never approaches the severety of emphasis placed on Japan's university entrance exam. However, the West has not been without victims from

examination anxiety. Nearly half century ago Brown (1938a, 1938b) called attention to the seriousness of the problem of test anxiety and in commenting on the causes of two students suicides at the University of Chicago, he noted:

> "One of these was definitely due to worry over an approaching exam and the other presumably was... These incidents show that students are taking their exams more and more seriously, and that the emotional reactions of students before exams is an important problem." (1938a, p. 11-12)

The recent trend at this university towards a decrease in the size of final exams contribution towards final grades is bound to have been welcomed by those who become apprehensive (test anxious) during exams. In spite of these changes many students are still so disturbed by test anxiety that they must seek professional assistance to help them cope with its debilitating effects.

Recounting his experience over a number of years at Duke University (USA), Charles Spielberger (1966b) noted that the number of students seen at university Pyschiatric Out-Patient Clinic "always seemed to increase during and immediately following university exam periods" and "anxiety concerning academic performance was in most cases either the salient symptom or an important background factor" (p. 361). These students complained that anxiety reduced their effectiveness while studying for and their performance during exams. Many reported that although they knew the answer to test questions they were unable to reproduce them because they "blocked" or "chocked-up" in the exam room. The result of all this was that "the students level of achievement was not commensurate with his intellectual aptitude, and his confidence in his own

abilities was seriously undermined" (Spielberger, 1966b, p. 362).

Although treatment of test anxious individuals had been going on much earlier the first test anxiety treatment study was published about twenty years ago (Spielberger & Weitz, 1964). The main aim of test anxiety treatment studies has been to decrease test anxiety, however, many of them including this one also aimed at improving cognitive-intellectual performance. As we will see in Chapter II the former aim has been far more easily and reliably obtained than the latter.

The purpose of this study was two-fold one was to see if a new treatment package (cognitive-attentional training plus study and examination skills training) is a more effective form of test anxiety treatment than others; the other was to identify personality variables that moderate the effectiveness of treatment and control for such factors thereby providing a more powerful and sensitive evaluation framework. The effects of personality variables on the effectiveness of test anxiety treatment has so far received scant attention and relatively rudimentary treatment.

Chapter I which follows this introduction deals with the nature of test anxiety and its theoretical formulations. In Chapter II a review of test anxiety treatment studies is presented together with other literature which pertains to the rationale behind the target treatment. Chapter III presents the rationale behind the other aspects of this investigation and Chapter IV describes the various measures employed. Chapter V describes how the investigation was carried out. In Chapter VI the results are reported and

and discussed in Chapter VII. The final section contains a brief overview and conclusions.

CHAPTER I

TEST ANXIETY AND THE TEST ANXIOUS

As Irwin Sarason (1980) a major contributor to this field, has pointed out in the preface of the only book entirely devoted to the topic, there is more to "test anxiety ... than meets the eye", and suggested that at times even workers in the field underestimate its scope. To this point he recounts that when he invited Donald Meichenbaum to contribute a chapter to the book (Meichenbaum & Butler, 1980) he replied that "he would try, but he wasn't sure there was either enough material or that he had that much to say". In the end this chapter was longer than they had agreed it would be but when Sarason tactfully pointed out this fact to Meichenbaum he said: "I can't cut out anything. If I do the whole thing won't make sense." (I. Sarason, 1980, p.ix). Perhaps related to the fact that it is such a global concept and difficult to come to grips with, test anxiety has probably been the single largest researched topic in pyschology and the literature relevant to it is likewise.

1.1 Definition

In a chapter of the book mentioned above, Sieber (1980) has pointed out the difficulties associated with the definition of test anxiety because it has important experiential components. Her discussion pointed out the shortcomings of (test) anxiety definitions offered in the literature. Levitt (1968), in his book "The Pyschology of

Anxiety", argued extensively on the difficulties associated with defining the construct of anxiety.

It is not without hesitation, therefore, that I shall try to define it : test anxiety is what test anxiety scales/questionnaires measure. Although this definition may appear to be a tautology and would anarguably be so some 30 years ago, three and half decades of research have validated the concept of test anxiety as what test anxiety indexes measure.

This is obviously a very pragmatic definition but undoubtedly very useful. Wittingly or unwittingly it has been the tacit definition of most contributors to the field . If there is "test anxiety" that falls outside the definition just given then it must fall outside the scope of this thesis and outside the scope of most if not all literature on the topic.

1.2 Nature of Test Anxiety

Although test anxious individuals have been differentiated along a number of dimensions, there are two lines of research which have been more extensively investigated thereby establishing more conclusively the nature of test anxiety along these two dimensions: (1) test anxious individuals tend to be more self-deprecatory, more self-preoccupied, and generally have a negative opinion of themselves; and (2) highly test anxious persons tend to perform less well on cognitive-intellectual tasks when compared with their low counterparts.

1.2.1 Test anxious cognitions

I. Sarason (1960) reviewed the literature on paper-and-pencil anxiety scales and pointed out that a number of studies provided evidence that highly anxious individuals are "more self-deprecatory, more self-preoccupied, and generally less content with themselves than subjects lower in the distribution of anxiety scales" [p. 404]. This applied to both individuals who scored high on test anxiety as well as on general anxiety measures.

The above conclusion was made on the basis of studies which reported relationships between anxiety scales and other paper-and-pencil personality measures. Along this line of research, Many and Many (1975), using a very large sample of pupils, found that self-esteem correlated -.38 and -.28 respectively with the Test Anxiety Scale for Children (TASC) and the General Anxiety Scale for Children (GASC) (S. Sarason, Davidson, Lighthall, Waite, & Ruebush, 1960), while Suinn and Hill (1964) found a much stronger relationship between self acceptance and the Test Anxiety Questionnaire (S. Sarason & Mandler, 1952) or the Manifest Anxiety Scale (MAS) (Taylor, 1953) in a sample of university students (-.58 and -.68, respectively). Goldfried and Sobocinski (1975) found high correlations between test anxiety and the Irrational Beliefs Test, which suggests that the ontogenesis of the highly test anxious negative preoccupations lies in their cognitive structures.

An interesting study was carried out by Nichols (1976),

who factor analysed the TASC and identified two factors: "poor self-evaluation" and "test anxiety". He then constructed a separate scale for each of them and found that the former correlated with the original scale more highly than the latter.

More direct evidence bearing on the fact that highly test anxious individuals tend to describe themselves in negative, self-devaluing terms and that they tend to emit self-disparaging cognitions has come from many studies that followed I. Sarason's (1960) review (Deffenbacher & Hazaleus, 1985; Galassi, Frierson, & Sharer, 1981a, 1981b; Galassi, Frierson, & Siegel, 1984; Heckhausen, 1982; Hollandsworth, Glazeski, Kirkland, Jones, & Van Norman, 1979; Holroyd, Westbrook, Wolf, & Badhorn, 1978; Houston, 1982; I. Sarason & Ganzer, 1962, 1963; I. Sarason & Koenig, 1965; I. Sarason & Stoops, 1978; S. Sarason et al., 1960).

The evidence also suggests that these types of self-derogatory and self-directed negative cognitions typical of the highly test anxious are triggered by their perception of the situation as stressful or evaluative. Deffenbacher (1978) gave a test to two groups of high test anxious subjects, one under low stress, the other under high stress conditions and similarly to two groups of low test anxious subjects and found that highly test anxious individuals reacted more negatively to the test and worried more than their low counterparts but only under stressful conditions. Ganzer (1968), Mandler and Watson (1966), Marlet and Watson (1968), Neale and Katahn (1968) as well as I. Sarason and Stoops (1978) reported evidence suggesting that the negative cognitions in the highly test anxious are triggered by the

stressful testing situation.

It is important to note that low test anxious individuals engage in negative and off-task cognitions also, however they do so only about half as much as their high counterparts (e.g., Hollandsworth et al., 1979; Mandler & Watson, 1966; I. Sarason & Stoops, 1978), moreover, Heckhausen (1982) has provided evidence suggesting that such cognitions are not as debilitating to low as they are to high test anxious individuals.

I shall end this section by noting three studies that are pertinent to the nature of test anxiety. Doris and Sarason (1955) provided bogus failure and success feedback alternatively upon completion of eight tests from various I.Q. scales. When questioned, high test anxiety subjects tended to blame themselves for the failures more than the low test anxious. I. Sarason and Ganzer (1962, 1963) asked their high and low test anxious subjects to describe themselves for half an hour during which certain classes of response were verbally reinforced. The most interesting finding was that highly test anxious subjects have a marked readiness to be conditioned when the response class being (positively) reinforced is negative self-references. However, when the (positively) reinforced response class is positive self-references they do not condition as strongly at all. These authors pointed out that Rogers (1960) using Taylor's MAS, a general anxiety index, did not find their highly anxious subjects to have a readiness to conditioning when reinforced for negative self-references, although Rogers like I. Sarason and Ganzer (1962) found that (positive) reinforcement increased the frequency of a given class of

verbal behaviour.

This situation where a given relationship is identified with test anxiety scales but not with general anxiety scales or it is identified more strongly and conclusively with test anxiety scales as opposed to general anxiety scales has occurred many times in the research which has investigated the relationship between anxiety and performance (see section 3.1).

1.2.2. Test anxiety and performance

A negative relationship between test anxiety and academic performance has been amply documented in the literature (Alpert & Haber, 1960; Boor, 1972, 1980; Bloch & Brackenridge, 1972; C. Brown, 1938a, 1938b; Carrier & Jewell, 1966; Daniels & Hewitt, 1978; Deffenbacher, 1980; Deffenbacher, Deitz, & Hazaleus, 1981; Dember, Nairne, & Miller, 1962; Doctor & Altman, 1969; Frost, 1969; Galassi, Frierson, & Sharer, 1981a; Galassi, Frierson & Siegel, 1984; Gaudry & Bradshaw, 1970; Gaudry & Fitzgerald, 1971; Harper, 1971; Marso, 1970; Milholland, 1964; Morris, Finkelstein, & Fisher, 1976; Morris & Liebert, 1970; Osterhouse, 1975; Pervin, 1967; I. Sarason, 1957b, 1959a, 1961b; S. Sarason et al., 1960; Sassenrath, 1967; Spielberger, Gonzalez, Taylor, Algaze, & Anton, 1978; Stanford, Dember, & Stanford, 1963; Walsh, Emgbretson, & O'Brien, 1968; Wittmaier, 1974) together with a negative relationship between test anxiety and measures of aptitude (Alpert & Haber, 1960; C. Brown, 1938b; Boor, 1972, 1980; Deffenbacher, 1977; Dember, Nairne, & Miller, 1962; Grooms & Endler, 1960; Morris & Perez, 1972;

Pervin, 1967; I. Sarason, 1957b, 1959a, 1961b, 1963; S. Sarason et al., 1960; Sassenrath, 1967; Walsh, Engbretson, & O'Brien, 1968; Walter, Denzler, & Sarason, 1964). The size of these correlations tends to be low to moderate. In general, test anxiety tends to correlate more strongly with aptitude measures than with academic performance. Global measures of academic performance like Grade Point Average (GPA) tend to correlate more highly and more consistently with test anxiety than performance on a given test or exam. Some studies failed to find a negative correlation between test anxiety and academic performance at least in some of the samples (Galassi, Frierson, & Sharer, 1981b; Grooms & Endler, 1960; I. Sarason, 1957b; Walter, Denzler, & Sarason, 1964) and in contrast with the rest of the literature S. Sarason and Mandler (1952) found a low but positive correlation between test anxiety and aptitude test performance.

The wide disparity in the size of the correlations obtained between test anxiety indexes and measures of academic and aptitude performance are probably due to two main factors. Firstly, test anxiety scales vary greatly among themselves and their predictive validity with respect with these two measures of performance vary accordingly (see section 3.1.3 for a critique of test anxiety measures). Secondly the circumstances surrounding the administration of a test anxiety measure substantially affect its relationship to academic and aptitude test performance.

Boor (1980) compared the correlation between performance on an IQ test as well as performance on a course exam with the Test Anxiety Scale (TAS) (I. Sarason, 1978) administered on two different occasions: straight after an exam and after

a regular class. The TAS correlated -.36 with exam performance when administered straight after it but only -.15 when administered after a regular class. The corresponding correlations with the IQ test were -.36 and -.20.

It is likely that because test anxiety correlates negatively and significantly with IQ test, it is not legitimate to partial out IQ from the relationship between test anxiety and academic performance; however, three studies have done exactely that. Daniels and Hewitt (1978) found that the zero-order negative correlation between academic performance and test anxiety was decreased but remained significant when IQ scores were partialled out. In two studies Boor (1972, 1980) found that the negative correlation between test anxiety and performance still remained negative, but it was no longer significant once IQ was partialled out.

1.2.2.1 Is High test anxiety always detrimental to performance?

Although the overall evidence leading to the conclusion that test anxiety negatively affects performance on cognitive-intellectual tasks is compelling, there is evidence that in individuals of very high ability test anxiety actually facilitates such performance (e.g., Gaudry & Fitzgerald, 1971). Although not directly relevant here because a measure of general anxiety was employed, Spielberger (1962) found that high anxious students in the top 20% of the ability distribution tended to do a little better academically than the low anxious students in the same ability range. In this and a study by Spielberger and Katzenmayer (1959), a negative correlation between general

anxiety and academic performance was obtained only for students in the broad middle range of aptitude.

Similarly, Paul and Eriksen (1964) found a negative relationship between test anxiety and "real life" exam performance only in the broad middle range of aptitude, which suggests that in their sample high test anxiety at least did not have a deleterious effect on performance for those subjects with the highest degree of aptitude. Denny (1966), Katahn (1966), and Gaudry and Spielberger (1970) found that high anxiety facilitates intellectual learning of subjects with high aptitude. Kight and Sassenrath (1966) found that high test anxiety facilitates performance of subjects high in achievement motivation.

Paulman and Kennelly (1982) found that subjects endowed with high test taking skills were not negatively affected by their high test anxiety on two aptitude measures if done separately. However, in the condition requiring them to complete both tasks concurrently high test anxiety did affect the subjects performance even among those with high test taking skills.

1.2.2.2. Test anxiety and task difficulty

There is considerable experimental evidence that the adverse effect of general anxiety on performance is greater on difficult tasks/items as opposed to easier ones (Carron, 1963; Katahn & Branham, 1968; O'Neil, Spielberger, & Hausen, 1969; Spielberger, 1966b; Spielberger & Smith, 1966). Ray, Katahn, and Snyder (1971) obtained the same finding using a measure of test anxiety. However, Ray et al., (1971) used a form of the State-Trait Anxiety Inventory (Spielberger,

Gorsuch, & Lushene, 1970) modified to apply to test anxiety while I. Sarason and Palola (1960), employing both a measure of general anxiety (MAS) and one of test anxiety (TAS), found much stronger results with respect to the TAS. Although the MAS yielded some significant relationships the TAS was the more strongly related to performance (The reader might recall the same situation occurring with the verbal conditioning studies: it is typical in this literature to find that test anxiety rather than general anxiety is the more crucial construct). I. Sarason and Palola (1960) found that when the task was easy, highly test anxious subjects tended to perform better than their low counterpart, but when the task was difficult the reverse happened.

Of more practical interest, Daniels and Hewitt (1978) analysed the effects of test anxiety on course exam questions varying in degrees of difficulty and found that the results obtained from experimental data were largely borne out: the influence of test anxiety was greater on difficult as opposed to easy items.

1.2.2.3. Test anxiety and instructions

Test anxious individuals are very responsive to social cues (e.g. Ganzer, 1968; Geen, 1976, 1977; Tobias, 1980). In this section, I shall briefly review studies that have investigated the effect of instructions upon the performance of high and low test anxious individuals.

When the situation is not defined as evaluative, highly test anxious individuals perform at about the same level as their low counterparts (e.g., Deffenbacher & Hazaleus, 1985; I. Sarason, 1972, 1975a, 1978; I. Sarason & Stoops, 1978).

However, when the situation is defined as evaluative and/or the experimenter gives "ego-involving" instructions, high test anxious individuals reliably do worse, while low test anxious ones reliably do better than under neutral, non-evaluative instructions (Deffenbacher, 1978; I. Sarason, 1956, 1957a, 1959b, 1961a, 1972; I. Sarason & Palola, 1960; I. Sarason & Stoops, 1978; Spence & Spence, 1966).

S. Sarason, Mandler, and Craighill (1952) administered an IQ subtest under two conditions. In one, subjects were expected to complete the subtest in a given period of time, while in the other they were not expected to complete in the same time period. The results showed that the high test anxious subjects did better in the latter condition while the low test anxious subjects did so in the former. Similarly, Morris and Liebert (1969) found that when timed, their high test anxious subjects did worse than their low counterparts on IQ subtests while the reverse was true when the same subtest were given untimed.

I. Sarason and Stoops (1978) found that highly test anxious individuals tend to overestimate the time waiting to begin an anagram task as well as the time spent working on the anagram, but only under acheivement-oriented instructions.

When the experimenter is reassuring, perhaps telling the subjects not to take the test "too seriously", "don't worry", "don't become too preoccupied", "take it easy", high test anxious individuals tend to perform better than the low test anxious (I. Sarason, 1958, 1972, 1978) and in any case, the high test anxious always do better under reassuring instructions than under control or non-evaluative conditions.

Paul and Eriksen (1964) compared students performance in a "real life" exam with performance on a similar exam where the experimenters were "as warm, permissive, and understanding as possible" and found that high test anxious students do better than their low counterpart on the latter conditions while the reverse was true under "real life" conditions. I. Sarason (1978) reported giving three of his undergraduate classes a no-risk second chance exam where the mark obtained on this exam would only count if it improved on the previous one. High test anxious individuals improved their performance on the second no-risk exam by 30% while the low test anxious improved only by 9%. I. Sarason himself pointed out the "scientific inelegancies" of this study; nonetheless, the data are highly suggestive.

Deffenbacher and Hazaleus (1985) failed to find that reassuring instructions facilitated the performance of highly test anxious individuals.

I shall quit this area of research by pointing out that in one study where the effect of reassurance and general anxiety as well as test anxiety was investigated (I. Sarason, 1958), test anxiety interacted with reassurance instructions. However, no such effect was obtained with respect to the general anxiety measure (MAS). Once again, test anxiety wins over general anxiety!

1.2.2.4 Effects of models

A number of studies have investigated the effects of modeling on the performance of high and low test anxious students.

Overall, the evidence clearly suggests that high test

nxious individuals are more strongly affected by the bservation of models. When the observed model displays :oping behaviour or otherwise succeeds in the task at hand ıigh test anxious subjects perform reliably better than under the no-model, control condition. On the other hand, the low test anxious subjects may or may not do so, but in any case their performance is not facilitated as much as the highly test anxious' (I. Sarason, 1972, 1975a, 1975b; I. Sarason, Pederson & Nyman, 1968).

Observing a non-coping model or a model who fails the task has a clear detrimental effect on the performance of the highly test anxious person but it works in the opposite direction with respect to the low counterpart (I. Sarason, 1975a, 1972).

I. Sarason (1968) investigated the effects of modelling upon the behaviour of high and low test anxious juvenile delinquent boys. There were two conditions; in one the high and low test anxious boys observed a model perform effective prosocial behaviour, while in the other, as well as observing the same model, each subject was replayed a videotape of the model and his own behaviour.

While exposure to the model produced positive results (in seven out of nine cases) in highly test anxious delinquents, contrary to expectation, when TV feedback was added to the modeling procedure, negative results where reported in six out of seven highly test anxious subjects. The results were not as strong but similar with respect to the low test anxious boys.

I. Sarason (1968) had suggested that the addition of the videotape feedback would enhance the effectiveness of the

modeling procedure; attempting to discover why his expectations were not borne out he interviewed his subjects and found that those in the modeling plus TV feedback conditions tended to attend to the discrepancy between themselves and the model thereby confirming their belief in their own inadequacies. Interestingly enough, such discrepancy did not bother subjects when both they and the models engaged in live role-playing.

1.2.3 Test anxiety and physiological arousal

Physiological arousal is regarded as one of the main components of anxiety (e.g., Spielberger, 1972a, 1972b) and test anxiety (e.g., Mandler & S. Sarason, 1952; Liebert & Morris, 1967). Moreover there is considerable evidence indicating heightened physiological activity as a result of the stresses of examinations on students (e.g., C. Brown & Gelder, 1938; Holroyd & Appel, 1980; Spielberger et al., 1978). In addition, highly test anxious students report higher levels of physiological arousal as shown by scores on test anxiety questionnaires that have a scale for worry and for emotionality (c.f. Deffenbacher, 1980).

Notwithstanding the above, studies that have tried to show a higher physiological arousal as measured by common pyschophysiological measures like skin conductance, heart rate, or respiratory rate, have failed to yield data that show highly test anxious individuals to be more physiologically aroused than their low counterparts (e.g., Hollandsworth, Glazesky, Kirkland, Jones, & Van Norman, 1978; Holroyd, Westbrook, & Badhorne, 1978; Holroyd & Appel, 1980).

Hodges (1976) reviewed studies attempting to show physiological correlates of state and trait anxiety (Cattel & Sheier, 1958; Spielberger, 1966) and found many relevant studies but very few of them reported significant correlates. Moreover, the results of these studies were mixed.

Galassi, Frierson and Sharer (1981a, 1981b) measured subjects' bodily sensations and in both studies these where found to be reliably higher among high than low test anxious subjects. However, significant associations between pyschophysiological measures and test anxiety remain the exception and not the rule. Holroyd et al. (1978) found highly test anxious subjects to show greater heart rate variability than their low counterpart, but they did not differ on either actual heart rate, spontaneous skin resistance response, or skin conductance level.

1.3 Test anxiety theory or dancing in doubt and fear?

Although interest in examination stress had begun in both Europe and the United States more than half a century ago (c.f. Spielberger et al., 1978), the study of test anxiety began in earnest some 35 years ago with Mandler and S. Sarason's (1952; S. Sarason & Mandler, 1952) now famous investigations on anxiety and learning.

These authors were the first to put forward a theory of test anxiety. Their theory was based on Hullian drive theory and assumed that the testing situation evokes two kinds of drives.

The first of these are learned task drives which are

reduced by "responses or response sequences which lead to the completion of the task" (Mandler & Sarason, 1952, p.166). The second kind is a learned anxiety drive, which can elicit two classes of responses: (a) those related to the task completion, which are anxiety reducing, and (b) those which interfere with task completion. According to this theory test anxiety is not necessarily debilitating; to be debilitating the individual must have learned that class of responses which interfere with task completion ((b) above). The Test Anxiety Questionnaire (TAQ) constructed by Mandler & Sarason (1952) was designed to measure this debilitating class of anxiety responses:

> These responses... may be manifested as feelings of inadequacy, helplessness, heightened somatic reaction, anticipation of punishment or loss of status and esteem, and implicit attempts at leaving the test situation. It may be said that these responses are self rather than task centred (p.166).

Unfortunately these authors did not concern themselves greatly with that class of anxiety responses which supposedly facilitates performance.

Alpert and Haber (1960) developed the Achievement Anxiety Test (AAT) which has a scale for facilitating anxiety (AAT+) and one for debilitating anxiety (AAT-). These authors did not dwell on the theoretical bases of their AAT and we are left to wonder whether it was the Mandler-Sarason theory of test anxiety or some other theory which suggested their two AAT scales or whether any theory at all suggested them. In any case the AAT- and the AAT+ consistently correlate negatively (c.f. section 4.1.1), which is not what the Mandler-Sarason theory would predict.

Based on factor analytical studies on TAQ items (Gorsuch, 1966; Sassenrath, 1964;) which identified essentially the same seven first order-factors and two second-order factors (emotionality and anxious avoidance of testing), Liebert and Morris (1967) conceptualized test anxiety as consisting of two major components, worry and emotionality, and developed a scale for measuring each of these components. Worry was described as "primarily cognitive concern about the consequences of failure" (p.975), while emotionality was described as consisting of autonomic reactions evoked by evaluation stress.

Although conceptualized as different components of test anxiety, worry and emotionality consistently correlate relatively highly (c.f. Deffenbacher, 1980). Worry consistently correlates negatively with performance while emotionality does on some occasions but not others (e.g., Deffenbacher, 1980; Morris & Liebert, 1969).

There is a strong descriptive feel about the Liebert and Morris theory but very little explanatory power. Richardson, O'Neil, Whitmore, and Judd (1977) factor analysed the TAS from a large sample of subjects (1200) and found that two factors accounted for 85% of the total variance; however the worry items and emotionality items loaded substantially on the same factor. Moreover, test anxious treatments designed to reduce worry (e.g., cognitive-attentional training) have been compared with treatments designed to reduce emotionality (e.g., desensitization), but no differrential decrease in either component of test anxiety was obtained by these treatments (Deffenbacher & Hahnloser, 1978; Hahnloser, 1974; Kirkland & Hollandsworth, 1980; McCordick, Kaplan, Smith, &

Finn, 1981; Osarchuck, 1974; Osterhouse, 1972; Scrivner, 1974; Vagg, 1974).

Deffenbacher and Parks (1979) and Snyder and Deffenbacher (1977) found that test anxiety treatments which would presumably reduce emotionality reduced worry rather than emotionality in an evaluative testing situation, while Deffenbacher, Mathis, and Michaels (1979) found test anxiety treatments designed to reduce emotionality reduced worry and emotionality equally.

I. Sarason (1960, 1972) has taken into account the situation as well as the personality of the highly test anxious individual in his studies and theorizing. When achievement is emphasised, high test anxious individuals perform more poorly than low test anxious individuals but when instructions are designed to reduce anxiety, high test anxious individuals show performance improvements while low test anxious individuals show performance impairments.

On the basis of extensive research evidence, I. Sarason (1972) concluded that high test anxious persons are more self-centred and self-critical than people who are low on test anxiety, and they are more likely to emit personalized derogatory responses that interfere with task performance. I. Sarason (1972) describes the behaviour of high and low test anxious individuals as follows:

> Whereas the less test anxious person plunges into a task when he thinks he is being evaluated, the highly test-anxious individual plunges inward. He either (1) neglects or misinterprets informational cues that may be readily available to him or (2) experiences attentional blocks. (p. 393)

I. Sarason's formulation is strongly based on past

research evidence but he stumbles trying to jump over research outlined in sections 1.2.2.1 and 1.2.2.2.

Wine's (1971, 1980, 1982) attentional approach to test anxiety is also based on past research evidence and has much of I. Sarason's flavour: (1) highly test anxious individuals are generally more self-preoccupied than the less anxious individuals and self-focused task-irrelevant cognitions are specifically elicited in high test anxious individuals when the situation is perceived as evaluative. (2) Test anxiety reduces the range of task cues used in cognitive-intellectual tasks. (3) Test anxiety is composed of cognitive and physiological components and it is the cognitive (worry) component that interferes most directly with cognitive performance. (4) The highly test anxious individual's cognitive task performance is improved by experimental and treatment manipulation designed to enhance attention to task relevant cues and reduce self-preoccupation and worry.

Wine's formulation is very similar to I. Sarason's probably because both are strongly anchored to past research evidence. However section (4) of this formulation is a questionable point. Although Wine (1982) cites a few treatment studies which support her point, a thorough examination of the relevant treatment literature (as presented in the following chapter) reveals that cognitive-attentional treatments (which aim at enhancing attention to the task and decreasing self-preoccupation and worry) seldom improve academic performance.

Test anxiety theory to date is a far cry from accounting for all the evidence that has accumulated on the topic. Any one of the test anxiety theories so far put forward crumbles

when confronted with data showing that highly test anxious individuals of high ability outperform low test anxious individuals, be they of high or low ability. Nor has test anxiety theory so far attempted to account for the evidence showing skill deficit associated with high test anxiety (Allen, Lerner, & Hinrichsen, 1972; Bruch, 1981; Culler & Holahan, 1980; Desiderato & Koskinen, 1969; Kirkland & Hollandsworth, 1979; Lin & McKeachie, 1970; Mitchell & Piatkowaska, 1974a; Wittmaier, 1972).

Test anxiety is a very rich and complex phenomenon and the theory that will adequately describe and account for all the available evidence may have to be equally complex. Test anxiety research overlaps with many fields of study and the theory that will adequately account for all its findings will cross many borders of psychological research.

CHAPTER II

TREATMENT RATIONALE, HYPOTHESIS, AND PREDICTIONS.

In this chapter I shall discuss the rationale, hypothesis and predictions with respect to four treatment procedures.

The chief aim of this investigation is to determine if the treatment package (PK) including Holroyd's (1976) cognitive-attentional training and a study skills training procedure with emphasis on examination-taking skills is effective in reducing anxiety and improving academic performance.

To assess whether the two components of this package work additively or synergically, the study also includes a treatment group to which Holroyd's cognitive-attentional training (CATH) is administered on its own, and one other which received study and examination skills (SES) only. A placebo (P) group was included to control for nonspecific treatment effects (Kazdin & Wilcoxin, 1976). In sum, PK is the target treatment while CATH, SES, and P function as various aspects of controls.

2.1 Test Anxiety Treatment : A Testing Arena?

As late as six years ago, Wine (1980) complained that a "cognitive-attentional interpretation of test anxiety has not permeated the treatment literature" (p. 373) and that such literature "continues to adopt an emotional reactivity interpretation" (p. 373) of test anxiety.

There are several possible reasons for this state of affairs. Firstly, the Wine (1971) and the I. Sarason (1972) cognitive-attentional theories of test anxiety are some 15 years old now and reflect the view that test anxiety is a predominantly cognitive phenomenon.

Secondly, the ideas elaborated by Wine and I. Sarason were present in much of the early test anxiety research beginning with the Mandler & Sarason (1952) study of anxiety and learning : the landmark study which provided the impetus in this area for research which has extended over three and a half decades and is devoid of any signs of decline.

Thirdly, the items of the most popular test anxiety indexes, for example the Test Anxiety Questionnaire (Mandler & S. Sarason, 1952), the Achievement Anxiety Scale (Alpert & Haber, 1960), the Worry-Emotionality Questionnaire (Liebert & Morris, 1967; Morris, Davis, & Hutchings, 1981) and the Test Anxiety Scale for Children (S. Sarason, Davidson, Lighthall, Waite, & Ruebush, 1960) reflect the notion that cognitions play a prime role in test anxiety. This connection between (1) test anxiety indexes, (2) the construct they were supposed to measure and (3) the way test anxiety was experienced (i.e., the face and content validity of these scales), although overwhelmingly cognitive, seems to have escaped the awareness of most test anxiety treatment researchers for about two decades.

Nonetheless, the prevailing use of desensitization and similar procedures in this treatment literature may not be a reflection of an emotional reactivity interpretation of test anxiety but rather an emotional reactivity interpretation of other forms of anxiety . This is a likely possibility and is

supported by statements and observations made by the authors of two thoughtful reviews of test anxiety research:

> These studies have evolved from interest in specific treatment techniques, rather than from an analysis of the nature and effects of test anxiety. Test anxiety has been considered to be a severe problem and thus its treatment is considered a stringent test of systematic desensitization procedures (Wine, 1971; p.101)
>
> Test anxiety has proven to be a usefl target construct for investigating basic behaviour change processes (Allen, Elias, & Zlotlow, 1980; p.155)

and

> In summary, our survey suggests that test anxiety continues to be exploited as a target for investigating the efficacy of a growing number of short term treatment investigations (Allen, Elias, & Zlotlow; 1980; p.165).

Are you beginning to forsee the founding of an organisation for the protection of the rights of test anxious students along the lines of the Anti-Vivisection Society?

Notwithstanding the above, the target treatment in the present investigation was devised on the basis of outcomes from treatment studies, correlational studies, and theoretical considerations. The aim of the PK treatment was to develop a procedure which would effectively decrease test anxiety to acceptable levels, improve grades, and possibly test performance as well.

2.2 Treatment of Test Anxiety

The conclusion derived from the first review of test anxiety treatment studies (Allen, 1972) was that they had been successful at decreasing self-reported test anxiety.

successful at decreasing self-reported test anxiety. However, the same could not be said with regard to performance improvements.

Several reviews that followed (Allen, Elias, & Zlotlow, 1980; Denney, 1980; Spielberger, Anton, & Bedell, 1976; Tryon, 1980; Wine, 1980) reached essentially the same conclusion as the first. Denney (1980), Tryon (1980) and Wine (1980, 1982) suggested that while the success rate at improving performance is not great (about 50% at best), the evidence is more encouraging when we consider treatment procedures designed to reduce self-preoccupation and worry and to enhance attention to task relevant cues (i.e., cognitive-attentional training techniques).

Nonetheless, the review that follows does not support, and is contrary to, the assertions made by both Denney (1980) and Wine (1980, 1982) as well as Tryon (1980), particularly when academic performance is considered.

2.3 On The Evaluation of Treatment

Typically a treatment has been evaluated with respect to its ability to decrease test anxiety and to increase academic and/or ability test performance. Moreover, these two criteria have been tacitly assumed to be along the same continuum. It will become clear by the end of this section why I believe this assumption to be misleading.

2.3.1. Criticism with respect to the test anxiety decrease criterion

The chief criticism I hold in this very important aspect of treatment literature is the researchers' seemingly blind faith in the data obtained through questionnaires.

A significant decrease in scores from a test anxiety questionnaire (relative to controls) has automatically been assumed to show that treatment was successful. Subjects are not asked whether they feel the treatment has helped them to manage their test anxiety and to what extent. The point I wish to make is that test anxiety scales have not been validated as measures of changing levels of test anxiety but merely as measures of test anxiety at one point in time.

A group of subjects may show significant decreases on a test anxiety index but we still don't know what their answer would be to the question: Did this treatment help you manage your test anxiety? In past research, it simply has not been asked. Furthermore, even in the face of a statistically significant decrease in scores of test anxiety measures, the clinical significance of a treatment remains to be ascertained. In short, the social validity (Wolf, 1978) of test anxiety treatment investigations has been disturbingly neglected.

Another important criticism I hold with respect to treatment evaluation is the overwhelming neglect of personality factors as possible moderators of treatment effectiveness.

Undoubtedly some subjects do not respond to treatment or

do not respond to the same extent as the group a whole, judging by the fact that there is a strong tendency for standard deviations to increase at post-test even though group means show significant decreases (c.f. Lent & Russell, 1978; Russell & Lent, 1982). Random assignment will not satisfactorily control the effects of these possible personality factors from moderating improvements differentially between groups because generally test anxiety treatment groups consist of only about 10 subjects and often much fewer.

The few studies that have investigated the effects of personality factors on the effectiveness of treatment (Mitchell & Ingham, 1970; McMillan & Osterhouse, 1972; Scrivner, 1974; Vagg, 1978) have employed simple ANOVA instead of more sophisticated and appropriate techniques like ANCOVA and MANCOVA (c.f. Cook & Campbell, 1979; Cooley & Lohnes, 1962; Overall & Klett, 1972). ANCOVA would not only indicate whether a particular personality variable acts as a significant treatment moderator, but also the extent to which it has a moderating effect.

2.3.2 Criticism with respect to the performance improvement criterion

Because of the negative correlation between test anxiety and academic as well as ability test performance (c.f. section 1.2.2) it has implicitly been assumed that a reduction in test anxiety would bring about an improvement in such measures of performance.

I believe this could be a misleading assumption to make, because the negative correlation between test anxiety and academic performance, for example, could be maintained by a constellation of other factors of which both test anxiety and academic performance are a function; as pointed out above test anxiety measures have been developed and validated as measures of test anxiety at one point in time.

Within the context of test anxiety and its relationship to academic performance, the question: is the test anxiety of a successfully treated and previously highly test anxious individual equivalent to that of an individual who all along has reported moderate or low levels of test anxiety?, cannot be answered confidently at this stage because other covariates, which could be affecting both test anxiety and academic performance, will not "allow" academic performance to increase as a result of test anxiety decreases.

In short, an important step has not been taken in the test anxiety treatment literature, namely: do test anxiety changes (decreases) correlate negatively with academic or ability test performance changes (increases) ? It is only when this question is answered affirmatively that we can expect a treatment designed to reduce test anxiety to also enhance academic and/or ability test performance.

Another important criticism relates to researchers and reviewers alike lumping performance on academic measures and performance on ability test measures under the rubric of "performance".

Apart from the fact that academic performance improvements are of much greater practical value than improvements on digit symbol or anagram tests (which have often been used in this

literature), the relationship between test anxiety and academic performance could be qualitatively different from the relationship between test anxiety and ability tests, even though the direction is the same.

Could we expect a successful test anxiety treatment to enhance academic performance but not ability test performance? We cannot confidently comment on this issue until we have investigated the nature of the relationship between test anxiety changes and changes in academic measures and ability tests measures.

Two more points of criticism before quitting this section, firstly: high test anxiety is not always detrimental and can enhance academic performance of those students with the highest academic ability. Therefore we cannot expect to achieve performance improvements among those students whose test anxiety does not seem to adversely affect their performance. What I am referring to is the research suggesting that test anxiety and academic performance correlate negatively only in the broad middle range of ability (c.f. section 1.2.2.1).

A situation could arise where one treatment group is composed entirely of subjects whose academic ability falls outside the broad middle range of academic ability distribution while a second treatment group receiving a different form of therapy is composed of subjects whose academic ability falls entirely within the broad middle range of the distribution. Now, suppose that both therapies are equally effective at decreasing test anxiety and that test anxiety decreases (regardless of treatment) result in academic performance increases only for subjects whose academic ability

falls in the broad middle range of the distribution.

The statistical analyses would show that both therapies are effective at decreasing test anxiety but only the second is effective in improving academic performance while in fact academic performance improvements were a result of the interaction between test anxiety treatment and subjects level of academic ability. The point is clear: we ought to exclude subjects who are not going to benefit academically from reduction in test anxiety when comparing treatment effectiveness in improving academic ability.

Finally, almost invariably researchers have drawn conclusions regarding the relative merits of the therapies they have investigated from the evidence of their own study and perhaps a few others. However, (1) because the sample sizes used are typically small, (2) because personality factors could play important moderating roles if not controlled for (as has largely been the case), and (3) because test anxiety is such a complex phenomenon, I believe that general conclusions based on a handful of studies are premature.

Considering the previous discussion, conclusions based on the merits of a particular form of therapy from one study alone are particularly hazardous. Nevertheless, many researchers have done so (e.g., Boutin & Tosi, 1983; Kirkland & Hollandsworth, 1980; Thyer, Papsdorf, Himble, McCann, Caldwell, & Wickert, 1981).

We shall see later that even if a few initial studies seem to suggest the superiority of a given form of treatment, many studies that follow may cast doubt on the initial conclusions. Before drawing general conclusions on the relative merits of

different types of treatment we need to look at the outcome of many studies, not just a few, or even worse, one.

The thrust of my argument is that because of the complexity of the test anxiety phenomenon we ought to put greater emphasis on replication as a strategy for validating the effectiveness of a test anxiety treatment procedure, as well as concentrating on eliminating threats to internal and external validity typical of quasi-experiments (Cook & Campbell, 1979).

Nevertheless, this discussion is not intended to detract from the methodological adequacy of the great majority of test anxiety treatment studies.

I believe the criticism made in this section is valid and important. The review that follows will rely on replication to assess the relative merits of treatment. It is reasoned that replication is equivalent to the use of large groups where personality factors would by random assignment balance each other across treatment groups.

2.4 Comparative Test Anxiety Treatment

This discussion concerns itself (unless otherwise specified) with studies carried out using university students as subjects, with the conclusions being drawn applying only to this population. We will see in a later section how techniques which have been successful with university students as subjects yield mixed and equivocal results when applied to school children (c.f. section 2.6).

A further clarification needs to be made here. The unit of measure in this review is the treatment group not the study. I believe it is important to take this step so that all instances of a particular technique being used can be accounted for, and also because some studies have used essentially the same technique twice but with slight variations. Several studies reviewed by Tryon (1980) under the "cognitive treatment" section employed a desensitization group for comparative purposes which were ignored under the "systematic desensitization" section. This problem can be avoided easily by using the treatment group and not the study as unit of measures; therefore comparative analyses are expected to be more valid and robust as a result.

Unless otherwise specified, groups receiving a treatment package (usually involving two separate procedures) were not included in the comparative analyses that follow for either form of treatment.

This step has been taken in order to exclude the possibility of giving credit to a given treatment when the effects could have been due to the two treatments acting synergically or additively.

Occasionally, when the writer believed one of the components of a treatment package to be virtually inert, only the seemingly active treatment was considered while the other was ignored for comparative analyses purposes; where this has been done it has been indicated.

It appears that just about all forms of behaviour modification techniques have been employed in the treatment of test anxiety, even if only to build evidence that a particular anxiety reduction technique is effective (as opposed to devising an effective treatment package for alleviating test anxiety).

Table 2.4.1 summarizes the outcome from treatment studies reviewed by Tryon (1980). Tryon (1980) did not distinguish in her review between treatment studies with university students as subjects and those with school children as subjects. Nonetheless a check of those studies yielding variant results indicated this was not a significant confounding factor.

Insert Table 2.4.1 here

Desensitization has been the single most used form of test anxiety treatment. Among the many studies reported by Tryon, 59 treatment groups received desensitization and of these, 50 (85%) reported a post-test decrease in self-reported test anxiety to a significant extent.

Academic performance improved in only 11 out of 27 (37%) groups. Performance on ability tests inproved on 6 out of 17 (35%) investigations. (Figures relating to performance

Table 2.4.1

Outcome of Test Anxiety Treatment Studies

Type of Treatment	Reduction in Self-Reported Anxiety		Improvement in Academic Performance		Improvement in Ability Tests	
	How many Successes (%)	How often Measured	How many Successes (%)	How often Measured	How many Successes (%)	How often Measured
Desensitisation	50(85%)	59	11(41%)	27	6(35%)	17
Self-Controlled Desensitisation/ Relaxation	16(94%)	17	4(80%)	5	4(50%)	8
Observational Learning	15(83%)	18	3(33%)	9	5(63%)	8
Cognitive Treatment	9(90%)	10	2(50%)	4	4(67%)	6
Study Skills	14(70%)	20	12(71%)	17	5(56%)	9
Total*	104(84%)	124	32(52%)	62	24(50%)	48

Adapted from Tryon (1980)

Note: *This is the sum of how many groups received a given technique and not the number of studies employing it, therefore the above total may exceed the total number of treatment studies reported by Tryon (1980)

improvements, both academic and ability tests, relate to those studies that measured it; many test anxiety treatment studies did not measure performance.)

Interestingly enough, the effectiveness of desensitization appears to be greater in those studies where this technique is the target treatment as opposed to those studies that used it for comparison purposes with different forms of (target) treatments.

Insert Table 2.4.2 here

We can see in Table 2.4.2 that when desensitization was used as a target treatment a significant reduction in self-reported anxiety was achieved in 88% of cases (36 out of 41), while it dropped to 78% (14 out of 18) when desensitization was used for comparison purposes.

Academic improvements were achieved in 42% (8 out of 19) of groups in the former case as opposed to 38% (3 out of 8) in the latter. With respect to ability test performance, desensitization as target treatment was successful in bringing about a significant improvement in 38% of the cases (3 out of 8), the corresponding figures for desensitization as comparison treatment was 33% (3 out of 9).

From the above it appears that, on the whole, desensitization is a more effective test anxiety technique when it is the sole target treatment than when it is not.

I would suggest that the effectiveness of a therapeutic technique is enhanced when the therapist himself or herself

Table 2.4.2

Outcome of Studies employing Desensitisation as Target and as Comparison Treatment in the Treatment of Test Anxiety

	Reduction in Self-Reported Anxiety		Improvement in Academic Performance		Improvement in Ability Tests Performance	
	How many Successes (%)	How often Measured	How many Successes (%)	How often Measured	How many Successes (%)	How often Measured
Desensitisation as Target Treatment	36(88%)	41	8(42%)	19	3(38%)	8
Desensitisation as Comparison Treatment	14(78%)	18	3(38%)	8	3(33%)	9
Desensitisation (Total)	50(85%)	59	11(41%)	27	6(35%)	17

Adapted from Tryon (1980)

believes that it is most effective. If a therapist believes that a given therapy is more effective he or she is likely to unwittingly convey his or her enthusiasm to the clients. This type of communication would probably be nonverbal; for example, the tone of the therapist's voice may be more convincing that a therapy is effective if they themselves are convinced that it is so. Steps aimed at avoiding such possible confounding were taken in the present study (c.f. sections 5.3 and 2.17) as well as checks to see that these efforts were successful (c.f. sections 4.7 and 2.17).

The evidence for the self-control form of desensitization or relaxation (i.e., with the addition of guided rehearsal which is introduced after the relaxation induction training) appears more encouraging than desensitazation. Tryon (1980) reported only 15 studies using this procedure and in these, 16 out of 17 (94%) treatment groups showed a reduction in self-reported anxiety and, of the five studies in which academic performance was measured, four (80%) produced a significant improvement. Moreover, of the eight treatment groups measuring performance on ability tests, significant improvement was obtained by four (50%). It is interesting to note the relatively small number of studies carried out using self-controlled desensitization/relaxation in spite of the fact that they appeared to be more effective than traditional desensitization.

Ten studies using observational learning (e.g., modelling and similar procedures) were located. Of the 18 treatment groups receiving an observational learning treatment, 15 (83%) showed a reduction in self-reported anxiety and 3 out of 9 (33%) showed an improvement in academic performance. In eight

of these studies, treatment effects were measured on ability tests and improvements were obtained on five of them (63%).

In addition to the above groupings Tryon (1980) had two more: cognitive procedures and study skills. In the cognitive procedures grouping she included 10 studies employing covert reinforcement, cognitive therapy, rational emotive therapy, and systematic rational restructuring.

Ten treatment groups received cognitive procedures. Of these, nine (90%) were effective in reducing self-reported test anxiety. Out of four that measured academic performance, two (50%) produced improvements. Four out of seven (57%) of the treatment groups that measured performance on ability tests showed improvements after treatment.

The study skills grouping included both studies that used study skills in combination with desensitization and studies that used study skills as a treatment on its own. It will be done likewise in this section.

In the eleven studies located, 20 treatment groups used study skills on its own or in combination (either with desensitization or relaxation); of these 16 (70%) produced a significant decrease in self-reported test anxiety. However, two studies (Lent & Russell, 1978; Mitchell & Ng, 1972) which included a treatment package as well as a study skills treatment on its own, found the former treatment to effectively decrease test anxiety but not the latter. Of the 17 treatment groups whose academic performance was measured, 12 (71%) reported improvements. With respect to ability test performance the success rate was 56% (5 out of 9). As mentioned above Tryon (1980) concluded that virtually

> "all treatments seem to be effective in reducing test anxiety ... Changing the student's academic performance is another matter" (p. 364).,

she adds,

> In looking at well designed studies that compare treatment and placebo groups, four procedures seem to change grades : study counselling (Allen, 1973), study counselling with systematic desensitization (Allen, 1971), study counselling with relaxation (Allen, 1973), and cognitive counselling (Holroyd, 1976). (p. 365).

Nonetheless, evidence collected in the present review casts doubt on some of Tryon's (1980) conclusions. "Study counselling", or what is referred to as study skills training (or treatment), is generally not effective at improving grades. Moreover, Harris and Johnson (1980), using the same study counselling procedure as Allen (1973), reported no significant improvement in grades. "Cognitive counselling" or what is referred to as cognitive coping techniques seldom produce improvements in grades.

The point I made above, that before drawing confident conclusions on the relative merits of a treatment one ought to look at the outcome of many studies using that particular form of treatment, remains valid. On the strength of this I hold reservations as to the effectiveness of "study counselling with relaxation" at improving grades. This is because the Allen (1973) study was the only one of its kind and, furthermore, the four treatment groups in this study consisted of two forms of study skills (group and self-administered) on their own and each of these treatments in addition to simple relaxation (i.e., without self-control instructions). All

four groups significantly increased in academic performance to an equivalent degree; in other words, study skills on its own was as effective as when in combination with relaxation in this study: The claim that "study counselling with relaxation" is in a general sense effective at improving grades remains to be established.

It should be pointed out that Mitchell, Hall, and Piatkowaska (1975) employed a treatment which involved study skills, academic counselling, and simple relaxation, but the design of this study is at variance with test anxiety treatment studies to such an extent that comparisons are not warranted. Moreover, this study was concerned largely with underachievement which appears to correlate with test anxiety (c.f. Mitchell & Piatkowaska, 1974a).

One might question why evidence from well controlled studies fail to be replicated. I believe the answer lies in subjects' individual differences which have largely been left uncontrolled. Apparent in test anxiety treatment literature is the implicit assumption that individual differences do not influence the effectiveness of a treatment or at least they are not important. This assumption may be misleading.

Tryon's (1980) optimism on the effectiveness of treatments packages, including desensitization and study skills, in improving grades was shared by other reviewers (Allen, 1972; Spielberger, Anton, & Bedell, 1976); moreover, with reference to grade improvements, Wine (1971, 1980, 1982) and Denney (1980) have been even more optimistic than Tryon on the merits of treatments designed to modify clients self-talk, beliefs, and attention. But let's look at the evidence.

2.5 Cognitive Coping Techniques

The ideas and assumptions behind these types of techniques have circulated for many centuries. A quote from Epictetus (1st century A.D.) says "It is not the things themselves which trouble us but the opinions which we have about these things." William James' (1884) theory of emotion hinges strongly on the idea that it is our perception of things which produces emotions.

So, it has long been recognized that an individual's perception of an event and his/her construction of the situation strongly affect the emotion that follows. However, it was Albert Ellis' (1962) contribution that made us appreciate the therapeutic significance of this fact, arguing that emotional disturbances that originate from illogical thinking can be remedied by teaching people to think more logically.

The rational emotive therapy (Ellis, 1962) method for altering irrational beliefs consists of a fairly direct verbal assault upon the client's thinking. However, Brehm (1966) has indicated through his concept of reactance that such an assault could easily provoke clients into a more tenacious position if they feel they are being coerced into changing their beliefs.

Denney (1980) has pointed out that the assimilation of the work of Ellis within the field of behaviour therapy has led to three major results.

Firstly, rational emotive therapy has been systematized into a more clearly operationally defined procedure (Goldfried, Decentesco, & Weinberg, 1974). Secondly, research

dealing with the impact of cognitive therapies has been conducted with much greater methodological sophistication. Thirdly, and in my view therapeutically most important, the types of cognitive coping techniques similar to the ones we are dealing with in this section have become more focused on challenging and changing specific negative and counterproductive self-statements clients emit in certain contexts, while emphasis on altering irrational beliefs has decreased.

Interestingly enough, focusing on narrower and more specific aspects of the problem has proved very fruitful in other areas of behaviour modification (for example, contingency management, relaxation).

The use of cognitive coping techniques has been prompted largely by Wine (1971, 1980, 1982) cognitive-attentional theory which has been discussed in the previous chapter. Briefly, test anxious individuals divide their attention between task-relevant and task-irrelevant thoughts while sitting a test. Task irrelevant thoughts, which are usually self-evaluative and self-deprecatory, have the effect of drawing attention away from the task at hand and detracting from test performance.

Within the context of test anxiety, cognitive coping techniques aim to identify client's task-irrelevant thoughts entertained during tests, eclipse such thoughts, and substitute positive self-statements which redirect their attention to the test.

Cognitive coping techniques have often been used in conjunction with desensitization or relaxation in an attempt to test if they have an additive or interactive effect on test

anxiety. Cognitive modification (Meichenbaum, 1972, 1977) involves the integration of cognitive-attentional training (CAT) and desensitization; this form of treatment is considered to be a combination not an addition (package) because desensitization is integrated within the cognitive coping rationale.

The outcome of these treatment studies can be found in Table 2.5.1. An examination of this table shows that 34 (97)% out of 35 such treatment groups were successful in bringing about a significant reduction in self-reported test anxiety.

Insert Table 2.5.1 here

With respect to academic performance improvements the results are not as impressive however; only 3 (20%) out of 15 treatment groups produced positive results. Performance improvements on ability tests were achieved by 11 (52%) of 21 treatment groups measured on this dependent variable.

It has escaped the attention of previous reviewers that the placebo groups employed by Allen (1971, 1973; c.f. Tables 2.7.1. and 2.8.1. respectively) were in effect very similar to cognitive coping techniques. Interestingly enough, both of these placebo groups significantly decreased self-reported anxiety and significantly improved academic performance.

This fact highlights the point that until we will have formulated a valid theory of test anxiety, devising a credible yet inert therapy (a placebo treatment) is a rather elusive task. By considering these two studies in addition to the

Table 2.5.1

Outcome of Studies Employing Cognitive Coping Techniques in the Treatment of Test Anxiety

Study	Conditions	No. of Treatment Sessions (duration minutes)	Weeks of follow-up	Reduction in Self-Reported Anxiety	Performance Improvement: Academic	Performance Improvement: Ability Tests
Barabasz & Barabasz (1981)	1. RET on 4 tapes 2. Study skills on 4 tapes 3. No Treatment	-	-	1>2=3	-	-
Bistline, Jaremko & Sobleman (1980)	1. Covert Modeling 2. Cognitive Restructuring 3. 1+2 combined 4. No Treatment	5(60)	8	1=2=3>4	1=2=3=4	-
Cooley & Spiegler (1980)	1. CAT 2. Relaxation 3. CAT + Relaxation 4. Placebo 5. No Treatment	5(55)	5	1=3>2=4=5	-	1=2=3=4=5
Crossley (1977)	1. CAT 2. Relaxation 3. Placebo	3(30)	-	1=2>3	-	1>2=3
D'Alelio & Murray (1981)	1. CAT (8 weeks) 2. CAT (4 weeks) 3. No Treatment	8(90) 4(90)	-	1=2>3	1=2=3	1=2=3
Deffen-bacher & Hahnloser (1981)	1. CAT 2. Relaxation 3. CAT + Relaxation 4. No Treatment	4(50)	5	3>2=1 4	1=2=3=4	1=2=3=4
Deffen-bacher & Hahnloser (1978)	1. Cognitive Restructuring 2. Relaxation 3. Cognitive Restructuring + Relaxation	?	?	1=2=3 pre>post	-	-
Dickinson (1983)	1. Cognitive Modification (Low Self-Control Ss) 2. Cognitive Modification (High Self-Control Ss) 3. No Treatment (Med. Self-Control Ss)	?	?	1=2>3	-	-
Fabick (1976)	1. Systematic Desensitisation 2. Cognitive Modification 3. Mantra Meditation	2(90)	-	3>2=1 1,2,3 pre>post	-	-
Finger & Galassi (1977)	1. CAT 2. Relaxation 3. CAT + Relaxation 4. No Treatment	8(45)	-	1=2=3>4	-	1=2=3=4
Goldfried, Lineham & Smith (1978)	1. Systematic Rational Restructuring 2. Prolonged Exposure 3. No Treatment	6(60)	6	1>2>3 1>2=3 (in testing situation)	-	-
Gordon (1983)*	1. CAT + Sublimal Stimulation (of "Sanctioned Oedipal...") 2. CAT + Sublimal Stimulation (neutral)	10(?)	-	1=2 pre>post	-	-
Hahnloser (1974)	1. Cognitive Restructuring 2. Relaxation 3. CAT + Relaxation 4. No Treatment	4(60)	-	3>2-1>4	1=2=3=4	1=2=3=4

Table 2.5.1 (cont.)

Outcome of Studies Employing Cognitive Coping Techniques in the Treatment of Test Anxiety

Study	Conditions	No. of Treatment Sessions (duration minutes)	Weeks of follow-up	Reduction in Self-Reported Anxiety	Performance Improvement: Academic	Performance Improvement: Ability Tests
Holroyd (1976)	1. CAT	7(60)	-	1=2=3=4>5	1>2=3=4>5	3=1>2=4=5
	2. Systematic Desensitisation					
	3. Cognitive Modification		4	1>2=3=4>5		
	4. Placebo					
	5. No Treatment					
Hussian & Laurence (1978)	1. Stress Inoculation Training (specific to test anxiety)	3(50)	-	1=2>3=4	1=2=3=4	-
	2. Stress Inoculation Training (general)		3	1=2>3=4		
	3. Placebo (discussion, some study skills)		35	1=2=3>4		
	4. No Treatment					
Kaplan, McCordick & Twitchell (1979)	1. CAT	10(60)	-	1>2=3>4	-	-
	2. Modified Desensitisation					
	3. Cognitive Modification					
	4. No Treatment					
Lavigne (1974)	1. Cognitive Modification	8(?)	-	1=2>3	1=2=3	1=2=3
	2. Systematic Desensitisation					
	3. No Treatment					
McCordick, Kaplan, Finn & Smith (1979)**	1. Cognitive Modification + 4	10(60)	-	3=2=1>4=5	3=1=2=4=5	1=2=3=4=5
	2. 1 + Videotape Modeling + 4					
	3. 1 + Rehearsal Modeling + 4					
	4. 40 min. Study Skills					
	5. No Treatment					
McCordick, Kaplan, Smith & Finn (1981)*	1. Cognitive Modification + 4	14(60)	-	3=2>4=5	1=2=3=4=5	1=2=3=4=5
	2. Pressured CAT + Modified Desensitisation + 4			1=4=5		
	3. CAT + Test Taking Practice + 4					
	4. 1 hr. Study Skills					
	5. No Treatment					
Meichenbaum (1972)	1. Cognitive Modification	8(60)	-	1>2>3	1>2>3	1=2>3
	2. Systematic Desensitisation					
	3. No Treatment		4	1>2>3	1>2>3	-
Osarchuck (1975)	1. Self-Controlled Desensitisation	6(60)	-	1=2=3=4	-	-
	2. Cognitive Restructuring					
	3. Cognitive Restructuring + Self-Controlled Desensitisation		8	1=2=3>4	-	1=2=3>4
	4. Placebo					
Reister, Stockton & Maultsby (1977)***	1. RET (Video + discussion + book)	5(105)	-	1=2>3	-	-
	2. Systematic Desensitisation					
	3. No Treatment					
Scrivner (1974)	1. Cognitive Modification	8(?)	-	1=2 pre>post	-	1=2 post>pre
	2. Systematic Desensitisation					
Vagg (1978)	1. Cognitive Coping					
	2. Biofeedback	7(60-90)	-	3=1>2=4	1=2=3=4	1=2=3=4
	3. Cognitive Coping + Biofeedback					
	4. No Treatment					

Table 2.5.1 (cont.)

Outcome of Studies Employing Cognitive Coping Techniques in the Treatment of Test Anxiety

Study	Conditions	No. of Treatment Sessions (duration minutes)	Weeks of follow-up	Reduction in Self-Reported Anxiety	Performance Improvement	
					Academic	Ability Tests
Wine (1971)	1. CAT 2. CAT + Relaxation 3. Placebo	6(60)	-	1=2>3	-	1=2>3
Wine (1974, Study I)	1. CAT 2. CAT + Relaxation 3. Placebo	6(60)	-	1=2>3	-	1=2>3
Wise & Haynes (1983)	1. Rational Restructuring 2. CAT 3. No Treatment	5(60)	- 35	1=2>3 1=2 (initial gains maintained)	-	1=2>3 1=2 (initial gains maintained)

Note: Between group significant differences are designated > ;
No significant differences are designated =.
* CAT was believed to be the only effective therapy and the other was ignored for comparison purposes.
** The Study Skills component in this study was considered to be too short to bring about a significant effect and therefore it was ignored for comparison purposes.
*** Reister, Stockton and Maultsby (1977), found one of their statistical analyses to indicate a treatment effect which only approached significance (p .07), nonetheless various other analyses indicated a significant treatment effect.

ones reported on Table 2.5.1, we find that the success of cognitive coping techniques at decreasing test anxiety is consolidated at 97% (36 out of 37) and climbs to 29% (5 out of 17) with respect to academic performance improvement.

As Allen (1971, 1973) did not measure performance ability tests, cognitive coping techniques in this respect remains at 52% (eleven out of twenty one).

It is evident that cognitive coping techniques are more effective in improving performance on ability tests than they are in improving academic performance: 52% versus 29%, respectively. This is interesting, particularly when we consider that most other forms of treatment appear to be more successful at improving the former than the latter (c.f. Table 2.4.1).

If we were to group both types of performance (academic and ability) tests, cognitive coping techniques would appear to be successful in 52% of the cases, as 13 of the 25 treatment groups which were measured on either or both types of performance produced improvements on at least one such measure.

I wish to point out again that if we aim to improve grades of highly test anxious students it is appropriate and useful to distinguish between academic and ability test performance and examine any pattern that appears, as has been done in this review.

Two additional studies using a cognitive coping technique deserve to be mentioned. They haven't been included in Table 2.5.1 for two reasons: (a) they have used a combination and an addition treatment respectively which are greatly at variance with the rest of the studies discussed in this section, and

(b) the design of these two studies is such that no conclusion can be made as to the merit of the cognitive coping techniques they have employed.

Boutin and Tosi (1983) use a true combination treatment, that is its components (cognitive restructuring and hypnosis) were integrated to produce a new form of treatment: rational stage-directed hypnotherapy. In the test anxiety treatment literature the term combination has been used where addition would have been more appropriate, as the component treatments were administered at separate times within a session and were not integrated with one another.

Boutin and Tosi (1983) found the combination treatment to be significantly more effective at decreasing self-reported anxiety than the hypnotherapy treatment alone. In turn hypnotherapy was significantly more effective than placebo which did not differ from no treatment. With respect to academic performance improvement the four treatments followed the same order. Unfortunately, cognitive restructuring treatment alone was not included in the design of this study and it is not possible to assess the relative merit of this component in the combination treatment, even if only with reference to this study.

Thyer, Papsdorf, Himble, McCann, Caldwell, and Wickert (1981) used a treatment which involved the addition of cognitive-attentional training, systematic desensitization, test taking practice, relaxation, and handwarming feedback. This treatment consisted of ten one-hour, six-weekly sessions and was compared to a treatment of equal duration which was a replication of the first plus the addition of distractions during the test taking practice. The rationale for the

distractions involved Wine's (1971) attentional theory: subjects would, in the face of such distractions, get more practice at directing their attention to the task. It appears that cognitive attentional training was carried out for approximately only one quarter of the total treatment time. Both groups reported significantly less test anxiety at post-test. The first treatment group reported less anxiety than the second, which was contrary to expectations, but the difference was not significant. Both groups improved significantly on performance measures, with the improvement being significantly greater for the first group on an anagram test but not on a manual dexterity test.

Noteworthy is the fact that subjects in this study were self referred or answered an advertisement, and were paid US$25. Himble, Thyer, Papsdorf, and Caldwell (1984) conducted a one year follow-up study on Thyer et al. (1981). Only fourteen of the initial nineteen subjects could be contacted. The only significant change at follow-ups was a further decrease on self-reported test anxiety.

These two studies appear very promising and it would be interesting to see if further investigations carrying out these types of treatment will confirm this initial optimism.

2.6 Cognitive Coping Techniques and School Children

Although most test anxiety treatment studies used university students as subjects, a number of studies have investigated the effectiveness of cognitive coping techniques on elementary and high school children. The outcomes of these studies are far less conclusive and far less promising than the evidence from university students suggests.

Of the five studies listed in Table 2.6.1, Dillard, Warrior, and Jaquilen-Perrin (1977), who employed a treatment which can be described as cognitive restructuring plus test-taking skills, made the only one that bears results which are consistent in themselves and with respect to the evidence reported in the previous section. These authors employed a very long treatment consisting of a 40-45 minute weekly session for six months. I shall return to this study in a later section (section 2.13).

Insert Table 2.6.1 here

Leal, Baxter, Martin, and Marx (1981) found cognitive modification to decrease self-reported test anxiety significantly. However they found systematic desensitization was not significantly different from no treatment controls. In spite of that, the desensitization group improved their performance on the Raven Matrices to a significant degree relative to the cognitive modification and the control groups, which were not significantly different from one another.

Table 2.6.1

Outcome of Studies using Cognitive Coping Techniques in the Treatment of Test Anxiety on School Children

Study	Conditions	No. of Treatment Sessions (duration minutes)	Weeks of Follow-up	Reduction in Self-Reported Anxiety	Performance Improvement	
					Academic	Ability Tests
Dillard, Warrior, Jaquilen-Perrin (1977)	1. Cognitive Restructuring + Test Taking Skills 2. No Treatment	25(40-45)	-	1>2	1>2	-
Leal, Baxter, Martin & Marx (1981)	1. Cognitive Modification 2. Systematic Desensitisation 3. No Treatment	6(60)	-	1>2=3	-	2>1=3
Little & Jackson (1974)	1. CAT 2. Relaxation 3. CAT + Relaxation 4. Placebo 5. No Treatment	6(60)	-	3>1=2=4=5	-	1=2=3=4>5 1=2=3=4=5 1>2=3=4=5 4>1=2=3=5
Maes & Heimann (1970)	1. RET 2. Systematic Desensitisation 3. Client Centered Therapy	10(?)	-	1,2,3 pre>post	-	-
Wine (1974, Study II)	1. CAT 2. Placebo 3. No Treatment	6(60)	-	1,2,3 pre>post	-	1=2=3 1>2=3

Note: Between group significant differences are designated >;
No significant differences are designated =.

This result is grossly at odds with test anxiety treatment literature results which suggest that (1) desensitization is effective in reducing self-reported test anxiety but relatively ineffective in improving performance on ability tests (c.f. Table 2.4.1); (2) there is a kind of "carry over effect: from test anxiety measures to performance measures. That is, when any of the treatment groups produce performance improvement, it would have reported reduced test anxiety. Moreover, group differences in performance improvements, had there been any, would be reflected by the differences in reported test anxiety (e.g., Allen, 1971, 1973; Boutin & Tosi, 1983; Harris & Johnson, 1980; Kirkland & Hollandsworth, 1980; Holroyd, 1976; Meichenbaum, 1972; Mitchell, Hall, & Piatkowaska, 1975).

The only exception to this "carry over effect" was Horne and Matson's (1977) study. Moreover, Horne and Matson's (1977) "maverik" treatment group received a study skills treatment lasting 10 hours: this result is not entirely unexpected in the light that study skills is the only form of treatment which is better at improving academic performance than it is at decreasing self-reported test anxiety (42% versus 38%; c.f. section 2.7). For all other forms of treatment the chances of decreasing self-reported anxiety are consistently greater than the chances of improving performance, the advantage varying from 12% (study skills plus desensitization package) to 68% (cognitive coping techniques) (c.f. Tables 2.5.1, 2.7.1, 2.8.1)

Little and Jackson (1974) also reported equivocal results which are inconsistent with the rest of the literature. While the addition of cognitive-attentional training and relaxation

training was effective in decreasing self-reported anxiety, neither of the two components differed significantly from the no treatment group nor from the placebo group.

With respect to performance and ability tests the results were again mixed. On the WISC similarities and on the DAT clerical speed and accuracy test all treatment groups, including the placebo group, improved significantly relative to controls, while on the WISC digit span test, WISC arithmetic test, and DAT verbal reasoning test there was no group difference. On the DAT numerical ability test the cognitive attentional training group improved significantly relative to the remaining four groups which did not differ significantly from one another. Probably the oddest result was on the DAT spelling test where the placebo group showed a significant improvement while none of the other four groups did.

Maes and Heimann (1970; in Rimm & Masters, 1974) measured only self-reported test anxiety and found that all three groups (rational emotive therapy, desensitization, and client centred therapy) reported significantly less test anxiety at post-test but they did not differ from one another.

Finally, Wine (1974, Study II) found that all three groups (CAT, placebo, and no treatment control) reported significantly less test anxiety at post-test and did not differ from one another. None of the three groups, however, improved on a reading test but the CAT group improved significantly on the IPAT test of G (an IQ test) while the remaining two did not and, further, were not different from one another.

The evidence from these few studies that have used

elementary and high school children as subjects bear equivocal results which differ considerably from the patterns set by treatment studies that have used university students as subjects. It would be interesting to investigate why school children respond to a given test anxiety treatment differently from university students. Nevertheless, the conclusion is inescapable that a treatment which is effective with a given population (university students) may not be so with a different population.

Noteworthy is the fact that the only study which seems to have followed the pattern of results obtained with university students as subjects (Dillard et al., 1977) involved a comparatively long treatment consisting of a 40-45 minute weekly session for a period of six months. We need more evidence to establish whether school children need longer time in therapy to gain similar benefits to those derived by university students from relatively short therapies.

2.7 Study Skills Training

The effectiveness of treatments employing study skills training on their own without the addition of desensitization will be considered in this section.

Much description as to the contents of study skills treatments employed in the literature is often lacking. However, when they are described it is evident that researchers tend to adopt the methods discussed by Robinson (1961, 1970) in "Effective Study". These usually include (a) a technique for textbook reading, (b) guidelines for efficient ways of using time, (c) advice on how to prepare for

examinations, and (d) techniques for tackling examination questions.

With respect to length and mode of presentation these treatments are a heterogeneous lot. Their duration has varied from 21 hours (Mitchell, Hall, & Piatkowaska, 1975) to 40 minutes (McCordick et al., 1979), with most studies employing about 6 to 8 hours of treatment. While most researchers have preferred a direct, face-to-face approach to study skills training (Allen, 1971, 1973; Lent & Russell, 1978; Mitchell et al., 1975; Mitchell & Ng, 1972) some have relied on videotapes (McCordick et al., 1979; 1981) and still others on audiotapes (Barabasz & Barabasz, 1981; D. Taylor, 1971).
Self-administered study skills have also been employed (Allen, 1973) and have both decreased self-reported test anxiety and improved academic performance to a significant level. Studies which have used a study skills treatment of less than two hours in duration (McCordick et al., 1979, 1981) were considered to be too short to be effective and have not been included in this analysis. Similarly, the Mitchell, Hall, & Piatkowaska (1975) study has not been included in this analysis because of its unusually long treatment (26 hours).

Insert Table 2.7.1 here

Table 2.7.1 lists seven studies that compared a study skills treatment with other formats of treatment. Six of these studies measured treatment effects on test anxiety and academic performance, while two of them (Bloom & Segal, 1977;

Table 2.7.1

Studies Employing Study Skills Treatments in the Treatment of Test Anxiety

Study	Conditions	No. of Treatment Sessions (duration minutes)	Weeks of Follow-up	Reduction in Self-Reported Anxiety	Performance Improvement: Academic	Performance Improvement: Ability Tests
Allen (1973)	1. Group Study Skills 2. Self-Administered Study Skills 3. 1 + Relaxation 4. 2 + Relaxation 5. Placebo (Attentional Training) 6. Testing Controls 7. No Contact Controls	7(60)	-	1=2=3=4=5>6	1=2=3=4>5>6>7	-
Bloom & Segal (1977)	1. Systematic Desensitisation 2. Study Skills	?	-	1>2	-	-
Cornish & Dilley (1973)	1. Study Skills 2. Systematic Desensitisation 3. Implosive Therapy 4. No Treatment	4(60)	-	2>3>4 2>1 3=1	1=2=3=4	-
Horne (1974)	1. Insight (Psychodynamic) 2. Desensitisation 3. Flooding 4. Study Skills 5. No Attention Control	6(?)	-	1=2=3=4=5	1=2=3=4=5	-
Horne & Matson (1977)	1. Modeling 2. Desensitisation 3. Flooding 4. Study Skills 5. No Treatment	10(60)	-	1=2=3>4=5	4>3 4>5	-
Melnick & Russell (1976)	1. Systematic Desensitisation 2. Hypnosis 3. Study Skills 4. No Treatment	4(45)	-	1>3=4 1,2 pre>post 2=3=4	1=2=3=4	-
Osterhouse (1972)	1. Desensitisation 2. Study Skills 3. No Treatment	6(60)	-	1>3 2=3	3>2 3=1	-
D. Taylor (1971)	1. Systematic Desensitisation 2. Study Skills (8 tapes) 3. No Treatment	8(40)	- 8	1>2>3 1>2>3	- -	- -

Note: Between group significant differences are designated >;
No significant difference is designated =.

D. Taylor, 1971) measured only test anxiety changes. Only three out of the nine (33%) study skills treatment groups reported a post-test reduction in self-reported test anxiety; while three out of seven (43%) produced academic performance improvements. None of the studies appearing on Table 2.7.1 were concerned with performance on ability tests. In addition to those listed in Table 2.7.1 there are six studies in Table 2.8.1 and one in Table 2.5.1 which employed this form of treatment. Grouping these latter studies with those on Table 2.7.1 we see that, of sixteen study skills treatment groups, only six (38%) produced self-reported test anxiety decrements and five out of twelve (42%) achieved an improvement on academic performance. Only two such treatment groups measured performance on ability tests but no improvement was apparent on this variable. In conclusion, compared to other forms of treatment, study skills is a poor technique for reducing test anxiety; however it compares evenly at improving academic performance. Not much can be said with respect to performance improvements on ability tests as they have been measured only twice; however the evidence so far is negative.

In passing, I should point out that the success of study skills training programs (of comparable content and length to those reported here) which aimed to improve academic performance of volunteer students not selected on the basis of high test anxiety seems to be about the same as those reported in Table 2.7.1 (c.f. Beneke & Harris, 1972; Jackson & Van Zoost, 1972; McReynolds & Church, 1973; Richards, 1975). Nevertheless, generally longer (between 7 hours to several semesters) study skills programs seem to bring about a significant academic performance gain in most cases (see

Entwistle (1960) for a review).

2.8 The Study Skills and Desensitization Package

As previously outlined, although in the literature these types of treatments are referred to as a combination, the writer believes this to be a misnomer as the study skills component is not combined with the desensitization procedure but merely added to it; in other words study skills and desensitization are administered at separate times within a given session and are not integrated with one another.

The study skills component has been described in the previous section. The desensitization procedure was based on Wolpe's (1958) reciprocal inhibition model and adapted to group therapy by Paul and Shannon (1966). Traditional (systematic) desensitization was subject to a number of variations (c.f. Harris & Johnson, 1980; Mitchell & Ng, 1972), however the basic tenet of learning deep muscle relaxation and learning to remain relaxed while mentally exposed to anxiety evoking stimuli was always retained.

The treatment package has usually been compared with study skills treatment only (e.g., Allen, 1971; Lent & Russell, 1978; Mitchell & Ng, 1972). Within a given study the length of therapy was kept constant across treatments. Obviously this strategy was adopted to control length and number of treatment sessions from acting as possible confounders. No researcher complained about the scarcity of time for therapy with respect to the two-component treatment and it was often remarked that even with less time available all items in the desensitization hierarchy were dealt with to completion (e.g., Holroyd, 1976).

The merits of this type of treatment in consistently decreasing test anxiety and improving academic performance have been pointed out by previous reviewers (Allen, 1972; Spielberger, Anton, & Bedell, 1976; Tryon, 1980). The present review points to the same conclusion: a treatment including study skills and desensitization is effective in decreasing test anxiety and improving academic performance.

In Table 2.8.1, there appear seven treatment groups that measured self-reported test anxiety and in every instance (100%) treatment was successful in bringing about a significant test anxiety decrement. Eight had academic performance measured and seven (88%) reported significant improvements. Only three such treatment groups were measured on ability tests performance and two of them (67%) achieved significant post-test improvements.

Insert Table 2.8.1 here

Mitchell, Hall, and Piatkowaska (1975) employed a treatment package which included study skills and desensitization, but the design of this study is at considerable variance with the studies considered in this section and in the rest of this review. Therefore it was not included in Table 2.8.1 nor was it used for comparisons made here.

This study, which I shall briefly describe, has employed treatments similar to the studies reviewed in this section.

Table 2.8.1

Outcome from Studies using Study Skills in Combination with Relaxation or Desensitisation

Study	Conditions	No. of Treatment Sessions (duration minutes)	Weeks of follow-up	Improvement in Self-Reported Anxiety	Performance Improvement: Academic	Performance Improvement: Ability Tests
Allen (1971)	1. Desensitisation	7(90)	-	3>1=2=4>5=6	3>1=2=4>5=6	-
	2. Study Skills					
	3. 1 + 2					
	4. Placebo (Attentional Training)					
	5. Testing Controls					
	6. No Contact Controls					
Allen & Desaulniers (1974)*			104	3=1=2=4=5=6 pre>post	3=1=2=4=5=6 post> pre	
Doctor, Aponte, Burry & Welch (1970)**	1. Study Skills + Systematic Desensitisation	11(?)	-	1>2 1,2 pre>post	1=2>3=4	
	2. Study Skills + Counselling	9(?)				
	3. No Treatment Volunteers					
	4. No Treatment Non-Volunteers					
Katahan, Stronger & Cherry (1966)	1. Study Skills + Systematic Desensitisation	8(60)	-	1>2	1>2	
	2. No Treatment					
Harris & Johnson (1980)	1. Individualised Covert Modeling + Study Skills	8(60)	-	1=2=3>4	1>2 1>4	-
	2. Self-Controlled Desensitisation + Study Skills					
	3. Study Skills					
	4. No Treatment					
Lent & Russell (1978)	1. Systematic Desensitisation + Study Skills	5(60)	-	1=2>3=4	1=2>4	1=2=3=4
	2. Cue-Controlled Relaxation + Study Skills					
	3. Study Skills					
	4. No Treatment		6	1=2>3	-	-
Mitchell & Ng (1972)	1. Study Skills Counselling					
	2. Desensitisation	9(50)	-	2=3=4>1=5	4=3>2=1=5	4=3>2=1=5
	3. Study Skills + Desensitisation					
	4. Serial Desensitisation + Study Skills		12	2=3=4>1=5		
	5. No Treatment		23	2=3=4>1=5		
McManus (1971)**	1. Desensitisation + Study Skills + Counselling	8(60)	-	-	1>2	-
	2. No Treatment					

Note: Between group significant differences are designated >;
No significant difference is designated =.
* This is a follow up to Allen (1971)
** The counselling component has been ignored for comparison purposes.

Mitchell et al., (1975) compared the effectiveness of three 26 hour treatments on test anxious failing university students (i.e., who had failed 3 or 4 out of 4 first year courses). The study involved five groups: (a) academic counselling (5 hrs) plus study skills (21 hrs), (b) academic counseling (5 hrs) plus relaxation (4 hrs), plus study skills (17 hrs); (c) academic counselling (5 hrs) plus relaxation (4 hrs) plus desensitization (9 hrs) plus study skills (8 hrs); (d) academic counseling (5 hrs); and (e) no treatment.

With respect to self-reported test anxiety reduction (c) was significantly more effective than (b) which in turn was significantly more effective than (a); (a) and (d) did not differ significantly and no anxiety data were reported for (e). On the academic performance criteria the same pattern was maintained: (c) was superior to (b) which was superior to (a); while (a), (d) and (e) showed no improvement. A follow up carried out 2 years later confirmed the superiority of treatment (a) over the other forms of treatment in terms of academic performance.

This study can be considered to be very successful considering that previous evidence indicated that "treatment procedures currently used to improve the academic performance of underachieving students have a low success rate (26 per cent); contribute little or nothing to either theory or practice; and are consequently a waste of valuable time and resources " (Mitchell & Piatkowaska, 1974a, p.19).

Before proceding it ought to be pointed out that a criterion for subject selection in several studies discussed in this section has been low study skills (Lent & Russell, 1978; Mitchell et al., 1975; Mitchell & Ng, 1972) or underachievement (Doctor et al., 1970; Mitchell et al., 1975) as well as high test anxiety. Furthermore, the only study that failed to bring about significant academic performance improvements did not use low study skills nor the underachievement criterion for subject selection (Harris & Johnson, 1980). Would the success of this type of treatment remain unchanged had this criterion for subject selection not been adopted? This is a question we cannot answer satisfactorily at this point in time.

A further study which should be discussed in this section is by Altmaier and Woodward (1981) who compared (a) a group receiving vicarious desensitization (i.e., observing a model undergo desensitization procedure) with (b) one receiving study skills, (c) one receiving a package involving the addition of both treatments and (d) one receiving no treatment. Treatment extended over six 50 minute sessions for the first two groups while the treatment package was twice as long (6 x 100 minute). Groups (a) and (c) reported significantly less test anxiety than (b) and (d), while no significant group difference ($p > .10$) was obtained with respect to academic performance improvements. Perhaps I ought to remind the reader that observational learning techniques have been found to be the least effective at improving academic performance (c.f. Tables 2.4.1, 2.5.1, 2.7.1, 2.8.1).

Before leaving this section I would like to note that Harris and Johnson (1980) found a group receiving a treatment

package involving the addition of Allen's (1971, 1973) study skills program and self-controlled desensitization to actually deteriorate on the academic performance criterion, though not to a significant level. This result is puzzling when we consider that (i) the same study skills program produced academic performance improvements in two separate studies (Allen, 1971, 1973), and (ii) self-controlled desensitization tended to succeed in studies employing this treatment which also measured academic performance gains (c.f., Deffenbacher, Mathis, & Micheals, 1979; Deffenbacher, Michaels, Michaels, & Daley, 1980; Denney & Rupert, 1977).

2.9 Maintenance of Gains

Allen (1972) in his review pointed out the lack of follow-up data. Allen, Elias, and Zlotlow (1980) complained that in "a disturbingly high proportion of studies ... participants simply seemed to disappear after post-treatment assessment was completed." On completing this review I must agree with Allen and associates. Nevertheless, a number of investigators have carried out follow-ups varying from as short as 3 weeks (Hussian & Lawrence, 1978) to 2 years after post-test (Allen & Desaulniers, 1974; Mitchell, Hall, & Piatkowaska, 1975). The results from following data suggest that the gains obtained as a result of treatment are largely maintained over time.

In the present investigation a follow-up was carried out four weeks after treatment and another twelve months after treatment.

2.10 A Case for Cognitive-Attentional Training plus Study Skills Treatment

The present review has investigated the validity of conclusions drawn by previous reviewers. The evidence suggests that cognitive coping techniques are an excellent form of treatment for reducing self-reported test anxiety, but the same cannot be said with respect to academic performance improvement as only five (29%) out of 17 such treatment groups have managed to show an improvement. Performance improvement on ability tests was reported in about half (52%) the cases, 11 out of 21.

Study skills training was effective at reducing self-reported test anxiety in 38% (6 out of 16) of the cases, and its chances of improving academic performance were 42% (5 out of 12). While a treatment package containing the addition of study skills and desensitization was shown to be the most effective form of test anxiety treatment with respect to both the anxiety reduction and the performance improvement criteria. It achieved 100% success rate by reducing self-reported anxiety every time it was employed, and 88% success rate at improving academic performance by increasing grades in seven out of eight instances.

We must bear in mind however that all of these latter studies which produced academic performance improvements had low study skills or underachievement as well as high test anxiety as criteria for subject selection.

On the face of this evidence two questions are begging for

an answer. Firstly, given that this form of treatment package has been recognized to be a most promising one since the beginning of test anxiety treatment reviews (Allen, 1972; Spielberger, Anton, & Bedell, 1976; Tryon, 1980), why have so few such treatment studies appeared in the literature? Secondly, how effective would a treatment package including study skills and cognitive-attentional training be? We can only speculate on the first question. Was it because, as Wine (1971) pointed out, test anxiety treatment studies have "evolved from an interest in specific techniques, rather than from an analysis of the nature and effects of test anxiety"? Or was it the beliefs of treatment researchers as to what test anxiety is that dictated the chosen form of therapy?

The second question needs to be investigated empirically. However, the evidence suggests that a test anxiety treatment package involving the addition of study skills and cognitive coping techniques would be effective and probably most effective in improving academic performance.

It appears that the effect of the two components of the treatment package including study skills and desensitization is additive and not synergyc or interactive. If we examine Table 2.8.1 we see that this treatment package has been compared directly with a study skills treatment six times; in five of these instances the former treatment has been significantly more effective at reducing self-reported test anxiety while no significant difference was found in the other. With respect to academic performance improvement the treatment package was found to be significantly more effective in three out of six occasions when such comparisons were made, while no significant difference was found in the remaining

three.

Again on examining Table 2.8.1 we see that two studies (Allen, 1971; Mitchell & Ng, 1972) compared this treatment package with desensitization alone. In these two studies three treatment group comparisons were made. The treatment package was found to be superior to desensitization at decreasing self-reported test anxiety in one such comparison but no significant difference was found in the remaining two. With respect to academic performance improvement the treatment package was found to be superior on all three occasions.

Although the number of studies available for this type of comparison is limited, two interesting points become evident from this analysis which suggest that the two components of this package are additive. Firstly, neither of the two components on its own has been found to be superior to their addition on any of these comparisons. Secondly, the outcome of these comparisons seems to reflect the relative strength of each component. (i) Desensitization has been found to be more effective at decreasing self-reported test anxiety than at improving academic performance, consequently the advantage of the treatment package over desensitization alone is expected to be greater with respect to academic performance improvement and not as great with respect to test anxiety decrements: the above comparisons show this to be the case. (ii) Study skills appears to be only a moderately effective form of test anxiety treatment, achieving a 38% success rate on the test anxiety decrement criterion, but it compares well with other forms of treatment on the academic performance criterion. Once more the expectation that the treatment

package advantage over study skills would be greater with respect to academic performance improvement has been confirmed by the above comparisons.

The other source of my confidence in the value of a treatment package which includes a cognitive coping technique and study skills stems from direct and indirect comparisons between cognitive coping techniques and desensitization. By adding the desensitization treatment groups reported by Tryon (1980) and summarized in Table 2.4.1 to those appearing in Tables 2.5.1, 2.7.1, and 2.8.1 and not reviewed by Tryon (1980) we find that desensitization has significantly decreased self-reported test anxiety in 56 (85%) out of 66 occasions, and significantly increased academic performance in 11 (38%) out of 29 cases and performance on ability tests in 9 (45%) out of 20 instances.

Let us now consider those studies that have been employed techniques which fall under the rubric of cognitive-attentional training (CAT); that is, techniques which aim at modifying two or more of the following: client's cognitions or self-talk, beliefs, and attention related to test taking. Rational emotive therapy has not been included because the writer believes it to be a less effective form of therapy than those that have evolved from it.

Cognitive modification and the addition treatment involving CAT plus relaxation or a form of desensitization are excluded from the present analysis for three reasons; firstly, by their very nature it would make a comparison with desensitization more difficult and less conclusive; secondly, these two types of cognitive coping techniques appear to be essentially as effective as CAT on its own (c.f., section

2.10); thirdly, the treatment package here argued for would save time and resources by eliminating the need to carry out desensitization or relaxation procedures without loss of therapeutic effectiveness.

An examination of Table 2.5.1 shows that it includes 20 treatment groups receiving CAT; in addition Allen (1971; 1973), in Tables 2.7.1 and 2.8.1 respectively, used as a placebo a treatment that essentially is a form of CAT. All of these 22 groups (100%) reported significant decreases in self-reported test anxiety, 3 out of 7 (43%) increased in academic performance, and 7 out of 12 (58%) increased in ability test performance. So we see that in this indirect comparison CAT is more effective than desensitization on all three criteria.

Unfortunately there have been very few direct comparisons, although in these comparisons CAT comes clearly first. Holroyd (1976) found CAT to be superior than desensitization on all three criteria. Kaplan, McCordick and Twichell (1979) found the insight component of cognitive modification (CAT) to decrease test anxiety significantly more than the desensitization component (modified desensitization) or cognitive modification itself. Unfortunately this study did not measure performance. Osarchuck (1975) reported no significant difference between CAT (cognitive restructuring) and self-controlled desensitization which both produced significant test anxiety reduction and significant ability test performance improvement. No measure of academic performance was taken.

The reader is reminded that the self-controlled form of desensitization or relaxation has been found to be generally

more effective than (traditional) desensitization (c.f. Table 2.4.1). Allen (1971) found no significant difference in test anxiety or academic performance between the desensitization and his placebo (i.e., CAT) groups.

In the limited number of comparisons availeble CAT is usually superior and definitely never comes second to desensitization. Clearly when the evidence is taken as a whole, it points to the conclusion that CAT is a more effective form of test anxiety treatment than desensitization. Will a treatment package which includes CAT and study skills be effective at decreasing test anxiety and improving performance (academic and ability tests) significantly? This is a question that ought to be answered empirically and it is the chief aim of this investigation.

2.11 Cognitive Modification and CAT plus Desensitization/Relaxation.

In this section I shall examine the evidence to find out how cognitive modification or CAT plus desensitization/relaxation compare with CAT on its own. These two treatments are very similar and differ only to the extent that cognitive modification involves some integration between CAT and desensitization, while there is no integration in the second treatment.

In Table 2.5.1 there appear 10 cognitive modification treatment groups, of these, 9 (90%) produced test anxiety decrements. Two out of five groups (40%) produced significant performance improvements on academic tasks and 4 out of 6

(67%) on ability tests. The respective outcomes for treatments which included CAT plus desensitization/relaxation were nine out of nine (100%) for test anxiety reduction, zero out of three (0%) for academic performance improvements and 3 out of 8 (38%) for ability tests performance. The success rate for CAT, with respect to the three criteria above, is 100% (all 22), 43% (3 out of 7) and 58% (7 out of 12).

The evidence points out that in this type of indirect comparison CAT and cognitive modification are roughly of equivalent effectiveness, while CAT is more effective than the package involving CAT plus desensitization or relaxation with respect to academic and ability tests performance measure. On the test anxiety reduction criteria both these procedures have the impeccable record of 100% success rate. Furthermore, it must be pointed out that about twice as many studies have employed CAT then either cognitive modification or CAT plus desensitization/relaxation; therefore we are more confident that the obtained success rate for CAT is closer to its true success rate while for the other two treatments more of their applications could reveal their success rate to change.

The number of available studies comparing these types of treatment with desensitization is very limited, as was the case for CAT. Cognitive modification was found to be significantly more effective than desensitization at decreasing self-reported test anxiety in one study (Michenbaum, 1972) while no such difference was found in four others; the same study found cognitive modification to be significantly more effective at improving academic performance but of equivalent effectiveness in two others (Holroyd, 1976; Scrivner, 1974). Holroyd (1976) found cognitive modification

significantly more effective than desensitization at improving peformance on ability tests, but the two treatments were found to be of equivalent effectiveness in three other studies. With reference to the other treatment of focus in this section, Osarchuck (1975) found a treatment package involving the addition of cognitive restructuring and self-controlled desensitization to be of equivalent effectiveness when compared to self-controlled desensitization on its own at both decreasing self-reported test anxiety and improving performance on the Wonderlic Personnel Test.

Once again, if we look in the previous section and compare the outcomes of these types of comparison with those made between CAT and desensitization we see that within the limited number of studies available CAT comes first. In other words the superiority of CAT over desensitization appears to be greater than that of cognitive modification over desensitization.

A number of studies have compared CAT with the two treatments of focus here. Two studies have compared CAT with cognitive modification and they point to the superiority of CAT. Holroyd (1976) found CAT to be significantly more effective than cognitive modification with respect to the test anxiety decrement criterion and the academic performance criterion; the two treatments were equally effective at improving performance on ability tests. Kaplan, McCordick, and Twichell (1979) measured only self-reported test anxiety and found CAT to be significantly more effective than cognitive modification.

CAT was found to be as effective as CAT plus desensitization or relaxation in six studies with respect to

the test anxiety decrement criterion but significantly less effective in two other studies. All comparisons involving measures of performance, academic (2) and ability tests (7), showed no significant difference between these two treatments.

Thus, in all of the above comparisons CAT is more effective than or as effective as cognitive modification or CAT plus desensitization/ relaxation, with the exception of two cases where this latter treatment was more effective than CAT at decreasing test anxiety. Nevertheless, in the light of the 100% success rate of CAT at decreasing test anxiety, these two exceptions appear of less importance.

As pointed out earlier, the number of studies available for these comparisons is limited; nevertheless the impression is strong that, as well as being superior in these comparisons, CAT is a somewhat more effective form of treatment when employed on its own than in addition or combination with desensitization or relaxation.

A further point I wish to make is that, given the limited number of studies available for the type of comparisons made in this and the previous section, Holroyd's (1976) study carries a great deal of weight in showing the superiority of CAT. Moreover Holroyd with his form of CAT has achieved the greatest group GPA increases reported in the test anxiety treatment literature (1,25). These are strong reasons for favouring the choice of CAT used by Holroyd (1976) (CATH).

2.12 Summary of CATH plus SES Package Rationale

There are several reasons which prompt the choice of CAT to be included in the treatment package proposed here.

(1). It appears that the treatment package which includes study skills and desensitization is an effective treatment for decreasing test anxiety and improving academic performance. Evidence concerning the effectiveness of this treatment in improving performance on ability tests is lacking. In any case, performance on ability tests in this literature has been more a matter of empirical and theoretical rather than of practical concern. Highly test anxious individuals are typically students (or pupils) and they are more concerned with improving their grades than achieving a higher score on the Wonderlic Personnel Test or on a Digit Symbol Test.

By making this point I don't intend to underestimate the practical value of improving an individual's score on ability tests which was previously impaired by high test anxiety. I am referring here to instances where personnel selection practices employ cut-off scores in one or more tests as prerequisite for successful job applications or for entry into a particular course.

(2). It appears that the two components (desensitization and study skills) of this package act additively and not synergically or interactively. Therefore, given that cognitive coping techniques are more effective than desensitization on all three evaluation criteria (test anxiety decrements, academic performance improvements, and ability test improvements), a package including a cognitive coping technique plus study skills is expected to be a superior test

anxiety treatment than the desensitization plus study skills.

(3). CAT appears to be as effective or more effective than other cognitive coping techniques which involve the inclusion of desensitization or relaxation; therefore, the choice of CAT seems to be appropriate on two accounts: seemingly greater effectiveness and saving of time and resources. (4). In the direct comparisons made earlier between CAT and its competing treatment, the CAT used by Holroyd (1976) (CATH) was superior to desensitization and cognitive modification. (5). Holroyd (1976), through his CATH, produced the greatest GPA increase (1.25) reported in this literature. (6). CATH is based on Wine's (1971, 1980) cognitive-attentional theory of test anxiety, which is a very popular and well accepted theory of test anxiety and has stimulated considerable research.

The inclusion of a study skills component with emphasis on examination taking skills (SES) has also a multi-faceted rationale.

(1). It appears that a psychological technique on its own is only modestly effective at improving grades.

(2). The addition of an educational technique involving training students to apply the guidelines developed by Robinson (1961, 1970), which itself appears of modest effectiveness at decreasing test anxiety and improving grades, seems to corroborate the effectiveness of desensitization at achieving these aims.

(3). Apart from the obvious expectation that improved study skills would improve academic performance (study skills correlate positively with academic performance, e.g., Brown & Holtzman, 1976; Cowell & Entwistle, 1971; Culler & Holahan, 1980; Desiderato & Koskinen, 1969; Entwistle

& Entwistle, 1970; Entwistle & Wilson, 1970; Jackson, Reid, & Croft, 1979; Kirkland & Hollandsworth, 1979; Lin & McKeachie, 1970; Mitchell & Piatkowaska, 1974a). There is ample evidence suggesting that study skills and test anxiety correlate negatively (Allen, Lerner, & Hinrichsen, 1972; Culler & Holahan, 1980; Desiderato & Koskinen, 1969; Kirkland & Hollandsworth, 1979; Wittmaier, 1972); therefore enhanced study skills is likely to both alleviate test anxiety and improve grades - the evidence collected in this review suggests this to be the case to a moderate extent.

A reservation I hold with reference to the applications of study skills training as a test anxiety treatement is that many researchers have not measured the level of study skills achieved as a result of a study skills treatment, therefore we are not in a position to establish whether the particular study skills training was effective at improving client's study skills and/or habits and further establish whether or not it is the improvement of such skills which may produce test anxiety reductions and or academic performance improvements. This is a step that the present investigation is set to take.

(4). The emphasis on examination taking skills has been prompted on three accounts: (a) There is evidence which suggests that test-taking skills correlate negatively with test anxiety and positively with academic performance (Bruch, 1981; Kirkland & Hollandsworth, 1979), therefore enhanced test-taking skills is likely to reduce test anxiety and improve grades. (b) I believe that a low score on a scale measuring test or exam taking skills is in practice, to a large extent symptomatic of uncertainty as to the correct

strategy to adopt when tackling a given exam question rather than the individual's erroneous beliefs as to the appropriation of his or her test-taking strategies. Teaching highly test anxious individuals to apply good test taking skills would therefore enhance task clarity and prevent task ambiguity to acts as a source of stress which is likely to make him or her anxious during exams or tests. (c) By perfecting their test taking strategies highly test anxious individuals are likely to direct their attention to the test (exam) itself at the expense of task-irrelevant rumination and worry. In other words, it is reasoned that the more natural and automatic tackling exam questions it becomes the less the likelihood that attention will shift away from the task at hand.

2.13 Results From Studies Employing Similar Treatments

Two studies have appeared in the literature employing treatments which are similar to the target treatment in this investigation. Dillard, Warrior, and Jaquilen-Perrin's (1977) subjects were 60 (29 males and 31 females) sixth-grade black children of low socio-economic status and, although as we have seen in a previous section (2.6) test anxiety treatments tend to produce mixed and equivocal results when administered to schoolchildren, this study is of interest because of the techniques employed.

Dillard et al., (1977) administered a treatment they called "applied test-taking self-control skills", which

appears to be essentially cognitive restructuring plus test taking skills with emphasis on enhancing subject's self-concept. Half the subjects received no treatment. Treated subjects showed a reduction in self-reported test anxiety and an increase in test grades. Interestingly enough, treatment appeared to have no effect on a self-concept scale. Furthermore, there was no sex difference on treatment outcome. It needs to be pointed out that this treatment, consisting of 40-45 minute weekly sessions over a period of six months, is very long when compared to the usual length of test anxiety treatments reported in the literature.

As previously pointed out, we need more evidence to establish whether schoolchildren need longer therapy time to gain similar benefits to that derived by university students from relatively short therapies.

Sixty highly test anxious university students were the subjects of Kirkland and Hollandsworth (1980), who compared the effectiveness of a treatment similar to the one employed by Dillard et al., (1977). The "skill acquisition" treatment of Kirkland and Hollandsworth consisted of training subjects to improve their test-taking skills, to learn to pay attention to the task rather than to task-irrelevant cues (attentional training), and to learn to evaluate themselves positively under testing conditions.

This treatment was compared to cue-controlled relaxation, meditation, and placebo (practice on anagram test only). Therapies consisted of 90-minute sessions. The three treatment groups reported significantly less test anxiety than placebo at post-test while no significant difference was evident between them. The "skill acquisition" group

significantly improved academic performance, anagram test performance, and a test-taking skills measure while no significant improvement was displayed by any of the other groups. None of the four groups improved on the Otis-Lennon mental ability test.

It is interesting to note that Kirkland and Hollandsworth (1980) have argued that test anxiety is a misnomer for what they see as being inefficient test taking skills, behavioural and cognitive. Accordingly, their "skill acquisition" target treatment was intended to enhance behavioural skills relating to test-taking and not as an anxiety reduction technique.

These two studies, despite their different rationale, are in many ways similar to the target treatment of this investigation. Their apparent effectiveness at both reducing test anxiety and improving academic performance enhances our confidence in the effectiveness of a CATH plus SES treatment package. Nonetheless, I believe a study skills component with emphasis on examination taking skills to be preferable to one which concentrates exclusively to test-taking skills.

Cited research shows highly test anxious individuals to have poor study skills and, on a long term basis, they are going to benefit more if they learn both test-taking skills and study skills.

In passing I should point out that McKordick et al., (1979) employed a 10 hour treatment which involved cognitive modification and a 40 minute videotape and found it significantly decreased test anxiety but failed to bring about any academic or ability test performance improvements. Moreover, McKordick et al (1981) employed a treatment which involved 13 hours of cognitive modification plus a one hour

study skills videotape and found it produced test anxiety decrements which only approached significance, while no improvements were achieved on either academic or ability tests measures.

The writer believes that it is unlikely that such a short study skills component would contribute significantly to treatment effectiveness. Moreover, the study skills only group did not differ from the no-treatment group on any of the measures in either study and a cognitive modification only group was not included; therefore we cannot even assess the contribution, if any, of the study skills component. In short, because of the treatment length, the inclusion of modified desensitization (as part of cognitive modification), and the short study skills component which the treatments employed, these two studies are not similar to the treatment package used in the present study.

2.14 How Did My CATH Differ From Holroyd's and Why.

After a few sessions it appeared that the subjects in the PK and particularly the CATH groups had learned quite satisfactorily the importance of replacing their own negative self-statements with positive ones, and were practising this skill. Consequently I felt that sessions were starting to become a little repetitive and in an effort to avoid this problem, and to further perfect the subjects' above skill, I introduced a role play exercise, some covert practice, and finally some direct practice on the Wonderlic Personnel Test. (I wish to point out that the practice test at the end of the SES treatment had been scheduled right from the start).

I think it is important to note that the role play exercise and the covert practice were introduced and designed to enhance subjects' skills at replacing their negative self-statments under testing conditions with positive ones and not as role play or covert practice at test-taking, thereby not altering the cognitive-attentional nature of CATH.

The practice test was given with the same intention and subjects were asked to practice the cognitive skills they had learned in the program. Furthermore, the subsequent discussion focused on the subjects' ability to employ these skills together with other suggestions from the therapist. The practice test was only 12 minutes long and it did not alter the cognitive nature of this treatment.

2.15 The CATH, SES, Placebo Treatments, and No Treatment Control Group

The CATH and the SES treatments have been included in this investigation to discover if the effectiveness of the package treatment is additive, synergic, or interactive thereby assessing if either component is redundant with respect to the criteria of effectiveness. The number of sessions and their length was planned to remain constant, thereby excluding therapy time from acting as a possible confound. Sessions were scheduled at an approximately similar hour, between 4 and 6 pm from Monday to Thursday, so that time of the day would also be excluded, as much as practical, from confounding the results.

The placebo (P) treatment was included to assess the degree of improvement that might be attributed to non-specific

treatment effects (Kazdin & Wilcoxon, 1976). Expectation of improvement; attention, warmth, and interest of therapist; and implicit demands for improvement generated by impressive therapeutic procedures are factors typical of all forms of treatment. Therefore, if we wish to disentangle them from the specific treatment effects of the target treatment we ought to control for them by the inclusion of a placebo.

The placebo treatment followed the same manual used by Holroyd (1976) for his placebo group. His "mind control" technique was chosen to function as placebo because it appears to be as credible as empirically effective theory and research based techniques (c.f., Holroyd, 1976). Nonetheless, Holroyd found this form of placebo to significantly reduce test anxiety and significantly improve academic performance but not ability test performance. The length of therapy and time of the day were as for the three treatment groups.

A no-treatment control group was not included in this investigation as repeatedly it has been found that non-treated subjects show neither self-reported test anxiety decrements nor performance improvement (c.f., Tables 2.5.1, 2.7.1, 2.8.1); therefore it is reasonable to assume that changes in the dependent variables are due to treatment effects, specific and non-specific.

2.16 Hypotheses and Predictions

The main hypothesis of this study is that the treatment package (PK) is effective at decreasing test anxiety and improving academic performance, and further that both CATH and SES are essential, additive components contributing to the

total effectiveness of the target treatment.

The second hypothesis is that non-specific treatment effects contribute to the effectiveness of PK as well as CATH and SES treatments, in other words, P will also produce positive results.

The degree of test anxiety decrease and performance improvements is predicted to be the greatest for the PK group, followed by the CATH, then by the SES, and P groups, in that order.

2.17 A Word on the Therapist

As pointed out in Chapter 5, I acted as therapist for all groups. The danger that knowledge of the hypothesis inherent to this study would affect the effectiveness of the various treatment was countered by efforts to keep the same level of enthusiasm, interest, and empathy in every session regardless of group. I believe I was successful in this effort.

I did not receive any formal clinical training and relied on my enthusiasm (which was great), reading of relevant literature, following the manuals, and talking to a professional therapist who had had extensive experience in treating highly test anxious students. To this point I wish to cite research, carried out by Allen et al., (1980) and by Di Tomasso (1981), arriving at the same conclusion using different techniques.

Both studies pointed out that therapist experience had no effect on treatment effectiveness. I wish to draw attention to the Hussian and Lawrence (1978) study where "an 18-year old female undergraduate psychology major" "trained by the first

author in the stress inoculation training procedure" served as one of the two therapists employed in that study, the other was a 23-year old male second year graduate student "whose clinical work was supervised by the second author" and "There were no systematic therapist effects" (Hussian & Lawrence, 1978; p. 31).

Moreover, Allen (1971) found two therapists to differ on a number of attributes as perceived by the subjects, but these differential collective subject perceptions were not related to variations in outcome.

The Participant Reaction Questionnaire (PRQ) administered at the end of treatment in this study assessed the subject's perception of the therapist on a number of relevant attributes and it is expected that it will detect any therapist biases towards the four groups.

It is predicted that the four treatment groups (including placebo) will rate the therapist and the treatments in equally favourable terms on all attributes listed on the PRQ.

CHAPTER III

RATIONALE, HYPOTHESES, AND PREDICTIONS RELATING TO OTHER ASPECTS OF THIS INVESTIGATION.

In this chapter I am going to discuss the rationale, hypotheses, and predictions related to the following aspects of this investigation: (a) test anxiety measuring instruments, (b) test anxiety self efficacy, (c) scholastic ability, (d) study and examination skills, (e) test and exam marks and final grades, (f) personality variables (g) social validity, and (h) follow-up data collection.

3.1 Test Anxiety Measuring Instruments.

Interest in and theorizing about anxiety dates back many decades (c.f. Spielberger, 1966a), but it wasn't until the 1930s and 1940s that important experimental work was done to study the problem of anxiety (I. Sarason, 1980).

Researchers took another step forward in their attempts to assess anxiety quantitatively and, although some work along these lines had been done more than a decade before (e.g., C. Brown, 1938a, 1938b), the 1950s saw a flowering of anxiety questionnaires, scales, and measures (I. Sarason, 1960). This flowering of anxiety indexes was concomitant with a dramatic increase in anxiety studies, both in terms of number and percentage of published studies in pyschology (c.f. Spielberger, 1966a)

3.1.1 General anxiety or specific anxiety?

In common with much of pyschology divergent positions emerged. Of relevance to this section are two groups of scholars who differed in their opinion as to how to best measure anxiety.

Underlying Janet Taylor's (1953) Manifest Anxiety Scale (MAS) is the assumption that there is a constant "level of internal anxiety or emotionality" and also "that the intensity of this anxiety could be ascertained by a paper-and pencil test consisting of items describing what have been called overt or manifest symptoms of this state" (J. Taylor, 1953, p. 285). Implicit in this position is the idea that there is a certain degree of anxiety characteristic of a personality which can be triggered by any stressful situation. Also holding this position were Kenneth Spence and others at Iowa Univerity (the Iowa group).

George Mandler and Seymour Sarason from Yale University, together with Richard Alpert, Ralph Haber and others (the Yale group) represented the alternative position which maintained that items concerning an anxiety measuring instrument should relate to the specific situations in which it is to be used (e.g., when measuring anxiety related to test taking, a scale should be used which contained items dealing only with test taking).

The Test Anxiety Questionnaire (TAQ) was constructed with items "specifically concerned with Ss attitudes and experiences in a testing situation" (Mandler and Sarason,

1952, p.166). Implicit in this approach is the idea that a given individual may experience anxiety in one stressful situation but not in other equally stressful ones, or at least that different but equally stressful situations will arouse varying degress of anxiety in the same individual.

Many studies have since provided support for the position held by the Yale group but not for that of the Iowa group.

Irwin Sarason, in a review of the literature relating to the use of anxiety scales, concluded that specific measures of test anxiety (e.g. the TAQ) tend to correlate negatively and significantly with performance while general measures of anxiety (e.g. the MAS) more often than not do not correlate significantly with performance. Nevertheless, Frost (1969) in his review concluded that evidence that test anxiety measures correlate more strongly than general anxiety is clearer for university students than it is for (elementary and high) school children. Walter, Denzler and I. Sarason (1964), although not reviewed by Frost, had results pertaining to tenth grade children which are consistent with Frost's conclusions.

Further evidence showing that results corroborated by studies using university students as subjects are not replicated in schoolchildren is reported in the previous chapter (Section 2.6). This suggests that schoolchildren's test anxiety is a phenomenon with different characteristics than university students' test anxiety. Many studies which have included measures of both test and general anxiety have consistently reported a correlation between measures of test anxiety and measures of performance stronger than that obtained using measures of general anxiety (Alpert & Haber,

1960; Carrier & Jewell, 1966; Desiderato & Koskinen, 1969; Grooms & Endler, 1960; Kirkland & Hollandsworth, 1979; I. Sarason, 1959a, 1961a; Sassenrath. 1967; Spielberger, Gonzalez, Taylor, Algaze, & Anton, 1978). Moreover, multiple regression analyses which included aptitude measures as well as test anxiety and general anxiety measures found aptitude to explain most of the variance in grade point average (GPA) (This result is typical when including intellective and non-intellective variables in a multiple regression analysis with academic performance as the dependent variable, c.f. Fishman and Panasella (1960)). However, test anxiety contributed to the prediction of this academic measure while general anxiety did not (Alpert & Haber, 1960; Kirkland & Hollandsworth, 1979; Sassenrath, 1967).

These results are even more significant when we consider that in these very three studies as well as in many others (e.g., Grooms & Endler, 1960; I. Sarason. 1957b, 1959a; 1960; 1961a), the correlation between academic aptitude and general anxiety was lower than the correlation between academic aptitude and test anxiety.

Consistent with the Yale group position, test anxiety measures intercorrelate more highly than general anxiety measures do (Alpert & Haber, 1960; Gordon & S. Sarason, 1955; Kirkland & Hollandsworth, 18979; Spielberger et al., 1978). Sassenrath, Kight, and Kaiser (1965) found that factors extracted from the TAQ related poorly, if at all, with factors extracted from general anxiety measures and concluded that there is no general factor "anxiety".

The results of Sassenrath et al.(1965) are even more striking when we consider that the TAQ is not as specific as

other test anxiety measures, e.g., the Achievement Anxiety Test (AAT) (Alpert & Haber, 1960). In fact the TAQ contains items relating to group IQ tests, to individual IQ tests, and course exams. Consistent with the rest of this literature Harper (1971) found that the TAQ section which deals with course exams correlates more strongly with GPA than the two sections which relate to IQ tests. Moreover, Sassenrath (1963) found that test anxiety affected concept learning while general anxiety did not. Sarason and Palola (1960) found test anxiety to affect test performance in three different studies while general anxiety was unrelated in two but was related, though less strongly than test anxiety, in the third.

If you thought that there was a match going on and were hoping for a come-back for general anxiety, I am sorry to tell you that the referee has just finished his ten count ... its a K.O.

Actually, support in favour of the Yale group and against the Iowa group was apparent even before these two positions emerged. Charles Brown (1938a) constructed an " examination neurosis" questionnaire with items dealing with university exams which correlated (-.22) with performance on examinations. He concluded that "students who became excited before examinations tend, on the whole, to do a little poorer in the examination than those students who are calm before the examination" (C. Brown, 1938b, p.30). Brown continued

> "The correlation obtained is small but highly suggestible. It must be remembered that the questionnaire used pertained to examinations in general; with the questionnaire revised so that it can be applied to one examination in particular we believe that the relationship between the questionnaire

> and achievement will be much higher" (p. 30-31)

Clearly, Charles Brown was a person with great vision.

Unfortunately C. Brown (1938a, 1938b) has been virtually forgotten in the mainstream of test anxiety research. Had it not been for World War II he would have probably received the credit he deserves. Here is another casualty never accounted for in war-time statistics.

It is evident from this short glide into the problem of measuring anxiety that research on test anxiety began about as early as research on general anxiety, and that test anxiety was the first identified specific anxiety (nowadays we recognize speech anxiety, interpersonal anxiety, heterosexual anxiety, social anxiety, and many more). Moreover, research on test anxiety has helped us greatly towards the understanding of anxiety. Nonetheless, researchers went a step further in the direction of specificity and formulated what can be called specific test anxiety.

3.1.2 Specific test anxiety

Liebert and Morris (1967) proposed the concept of two components of test anxiety, worry and emotionality, and developed an instrument (the Worry-Emotionality Questionnaire (WE-Q)) with a scale for each of them. Worry can be described as cognitive concern and fear of loss of self-esteem while emotionality is tantamount to perceived physiological-affective reactions and arousal. The Inventory of Test Anxiety (ITA) and the Test Anxiety Inventory (TAI),

which were developed by Osterhouse (1972, 1976) and Spielberger and associates (1978) respectively, also have a scale for worry and for emotionality. Morris, Davis and Hutchings (1981) have recently produced a revised form of the W-EQ.

The specificity of the W-EQ is enhanced by the fact that its items pertain to one examination or test to which the questionnaire is supposed to be administered before it begins and/or after it is finished.

Reviewing those studies which have investigated the relationship between worry and emotionality and academic performance, Deffenbacher (1980) drew two major conclusions. Firstly, worry consistently forms an inverse relationship with performance expectations and actual test performance while emotionality does not, forming a negative relationship in some cases but no relationship in others. Secondly, worry accounted for more performance variance than emotionality. Most importantly, when emotionality was partialled out worry maintained its direction and size of relationship with performance. On the other hand when worry was partialled out there remained no significant relationship between emotionality and performance and its direction now tended to be positive. More recent evidence (Deffenbacher & Hazaleus, 1985) is consistent with this finding.

3.1.3 A carousel of test anxiety questionnaires

Measurement is one aspect of the test anxiety literature which largely reflects the state of the latter, a large body of research carried out in a piecemeal fashion with little

evidence of cumulation.

What follows is a brief review of a variety of test anxiety measuring instruments and a short critique of this aspect of research. As this section is not intended to be an exhaustive review, only the more popular instruments are considered.

The first widely used test anxiety measure was the Mandler-Sarason Test Anxiety Questionnaire (TAQ). The Test Anxiety Scale (TAS) originally consisted of 21 true/false items rewritten from the TAQ (I. Sarason, 1958). I. Sarason and Ganzen (1962) later presented a 16-item TAS which correlates .93 (I. Sarason, Pederson and Nyman, 1968) with the current 37-item scale (I. Sarason, 1972; 1978; 1980). The Alpert-Haber Achievement Anxiety Test and the TAS have been the two most popular test anxiety research instruments to date. However a variety of different instruments have been devised and heavily used in this literature.

Some test anxiety questionnaires make reference to intelligence test taking as well as classroom exams (e.g., the TAQ and the TAS). Others refer to anxiety felt while studying for exams or tests as well as while actually taking them, for example the Suinn Test Anxiety Behaviour Scale (STABS) (Suinn, 1969).

The ITA deals only with anxiety experienced in a recently taken exam or test or one just taken, while the W-EQ can be used to measure the anxiety experienced before a test or during a test just sat, depending on the time of its administration.

The Cognitive Inference Questionnaire (CIQ) by I. Sarason (1978, 1980) is intended to be administered straight after a

cognitive-intellectual task and it measures the extent of task-irrelevant cognitions an individual has experienced during that task.

Several test anxiety instruments make no reference to a particular test but measure test anxiety as it is typically experienced, for example, the AAT, the TAS, the TAQ, the TAI. One could make an analogy here with the concept of a distinction between state test and trait test anxiety (Cattell & Scheier, 1966; Speilberger, 1966a.) Instruments like the ITA the W-EQ and the CIQ measure "state test anxiety". This is a transitory emotional state typical of the test or exam referred to by the testee but not necessarily indicaticating a tendancy to become anxious during tests or exams. On the other hand the AAT, the TAS, the TAQ and the TAI measure "trait test anxiety", that is a tendency to become anxious during any exam or test.

As previously pointed out, the W-EQ, the ITA and the TAI have a scale for worry and one for emotionality, whereas most of the other questionnaires have only one scale. The TAS and the CIQ predominantly measure worry while the STABS and the Mathematic Anxiety Rating Scale (MARS) (Richardson and Suinn, 1972; Suinn, Edie, Nicoletti, and Spinelly, 1972) predominantly measure emotionality. The MARS measures anxiety aroused by mathematics tests, and a variety of behaviours relating to computations and analytical problems.

The heterogeneity of test anxiety indexes is also apparent in their length: the MARS, the STABS, and the TAQ, with 98, 50, and 42 items respectively are long, while the CIQ has only 12 items and the WEQ only 10.

The AAT is unique in that it is the only test anxiety

instrument that considers one's affective reactions to evaluative situations in their debilitative as well as their facilitative functions. Accordingly it has a scale for facilitative test anxiety (AAT+) and one for debilitative test anxiety (AAT-). This questionnaire is described and discussed in the following chapter (section 4.1.1).

While all the above instruments have been designed with adults and young adults in mind, several test anxiety questionnaires have been constructed to measure test anxiety in children. The Test Anxiety Scale for Children (TASC) (S. Sarason, Davidson, Lighthall, Waite, and Ruebush, 1960) has been the most popular of them. Stanford, Dember, and Stanford (1963) have adapted the AAT into a children's form.

The 1950s flowering of anxiety scales was viewed by a few observers of the psycological scene with more than a little alarm, arguing that the garden was infested with too many weeds. (I. Sarason, 1980). But is the test anxiety "garden" infested with too many weeds? This is very difficult to establish at present, given that a systematic study attempting to assess the comparative validity and reliability of the various test anxiety instruments has so far not been undertaken. The garden has not been cleared of its weeds and its plants have not been adequately cultivated. Indeed, in the present state of affairs, all but the most thorough and careful test anxiety scholar cannot clearly distinguish weeds from plants in the garden, nor are there adequate instructions of when and how best to use the plants.

There is everything from wide disparity to close affinity among the various tst anxiety indexes, and quite clearly there is a need to establish how they vary among themselves and by

how much. Furthermore many researchers and reviewers discuss how test anxiety was significantly decreased in one group (more or less than another group) or how test anxiety is significantly related or unrelated to a particular variable or construct, but all too often fail to identify the instrument used to measure test anxiety.

It is hoped that the above discussion has pointed out the need to do so. Unfortunately, because it is beyond the scope of this thesis, especially when we consider that the literature relevant to it is very large, I have not and shall not try to identify test anxiety by its measuring instrument. Furthermore, again because of the large literature involved, I had and shall have to rely upon conclusions made by previous reviewers and researchers who have typically not identified test anxiety by its measuring instruments.

3.1.4 Why did you choose the AAT?

The need to use measures of test and not general anxiety is obvious from the discussion at the beginning of this section (3.1). Of course, both AAT scales measure anxiety experienced in testing situations and therefore it is an appropriate instrument for measuring changes in levels of anxiety following a test anxiety treatment programme.

The specificity of this instrument (its items deal with academic examinations) makes it appropriate for this investigation in which all subjects were university students who were troubled by anxiety experienced during university exams and tests, and wanted to overcome such anxiety.

The AAT is probably the best validated test anxiety measure (Albert & Haber, 1960; Dember, Nairne & MIller, 1962; Kirkland & Hollandsworth, 1979; McCordick et al., 1981; Milholland, 1964) and it comes first when validated alongside other popular and reliable test anxiety measures (Kirkland & Hollandsworth, 1979; McCordick et al., 1981).

One of the validation criteria of the AAT was an inverse relationship between AAT- and academic performance. With a high AAT- cut-off score as a prerequisite for participating in the treatment programme, it was likely that those students whose test anxiety adversely affects their exam performance would be selected.

I am implying here that with certain highly test anxious students (i.e., those in the top 20% of the academic ability distribution) exam performance is not debilitated by test anxiety as measured by some other test anxiety instruments. On the contrary, it may be facilitated by it (but see a discussion on this topic in section 1.2.2.1). There is considerable research supporting this fact (Denney, 1966; Gaudry and Fitzgerald,1971; Gaudry and Spielberger, 1970;1971; Katahn, 1966; Kight and Sassenrath, 1966; Spielberger 1962;1966b; Spielberger and Weitz, 1964).

Indeed I ought to point out that in the constructions of their AAT Alpert and Haber's main criterion for item inclusion was that it correlated highest with academic aptitude test performance. Some AAT- items would correlate inversely and relatively strongly with academic performance in the middle range of aptitude but would not correlate or would correlate positively in the top range of the distribution. Such items would possibly yield a low overall correlation with

performance, independent of aptitude, and therefore be excluded from the final AAT-. In view of this reasoning it is not surprising that Pervin (1967) did not find differential predictability of the AAT-, AAT+, or their difference at different levels of aptitude.

Recent research has pointed out that the AAT is an excellent test anxiety measure (Kirkland & Hollandsworth, 1979; McCordick et al., 1981) and I believe it is possibly the best to date. Furthermore, because the AAT is a very popular instrument in the test anxiety treatment literature its choice allows for more direct and ready comparisons with many similar studies.

Huck and Jacko (1974) pointed out that the difference between AAT- and AAT+ (AAT-+) is a better measure of test anxiety than either scale on its own; Kirkland and Hollandsworth (1979) findings corroborated this finding. One of Finger and Galassi's (1977) as well as Kirkland and Hollandsworth's (1980) subject selection criteria for their test anxiety treatment study was an AAT-+ score falling in the top third and 28% of this distribution of scores respectively.

Although I was unaware of these studies at the time this investigation took place I believe that even though the AAT- and AAT+ difference is a better measure of test anxiety obtaining higher reliability estimates (Huck & Jacko; 1974) and correlating more strongly with aptitude as well as academic measures (Kirkland & Hollandsworth, 1979) than either scale on its own, it is doubtful whether a cut-off score on the AAT-+ is more appropriate than a cut-off score on the AAT-, as has been used in this investigaion and many others.

The reason why I hold such doubt is that test anxiety

treatments have rarely succeeded in significantly raising group AAT+ while they have usually effectively decreased AAT- significantly. Now, a given cut-off score on the AAT-+ can be a result of very low AAT+ and moderate AAT- rather than a very high AAT- and low or moderately high AAT+. Given that test anxiety treatments are generally ineffective in reducing AAT+ but generally effective in reducing AAT- we can see that in the first instance (moderate AAT- minus very low AAT+) the cut-off score on the difference scale is not appropriate. A student with a moderate level of AAT- is not going to benefit greatly from a further decrease, moreover it is not known whether test anxiety treatments are effective at reducing debilitative anxiety of students with low or moderate test anxiety levels. In any case, we run the risk of failing to provide treatment for those who need it most. Had I known of the above studies at the time of subject selection I would still have used the same debilitative cut-off score because I believe it is likely to select those students who will benefit most from treatment. In short, the returns (in terms of test anxiety alleviation) are likely to be greater when we use students who score in the top 13% of the AAT- distribution than when we use students who score in the top 13% of the AAT- minus AAT+ distribution.

I employed an AAT- score of 32 or greater as a subject selection criterion. Holroyd (1976) found such a cut-off score to include the top 13% of the distribution on this measure. He also used a cut-off score of 32 on the AAT- as a subject selection criterion for his treatment investigtion. An even higher AAT- cut off score would have been more desirable in order to give treatment to those who need it

ıost, but one runs the risk of not finding enough subjects if .he subject selection criteria are too restrictive. Greater lescription and more information on the AAT can be found in :he next chapter.

3.1.5 ... And the ITA ?

On the basis of the discussion in section 3.1.2 I felt it was both important and interesting to measure changes in worry and emotionality. Therefore I had to include a questionnaire with scales for worry and for emotionality.

The choice was between the W-EQ, the ITA, and the TAI, because the AAT is a measure of "trait test anxiety". Because I preferred a measure of "state test anxiety" the TAI was excluded leaving W-EQ and the ITA. These two questionnaires are very similar except that the ITA is longer than the W-EQ.

Osterhouse (1972) found a split-half reliability (corrected for length) of .92 for the ITA, while Morris and Liebert (1970) found alpha coefficients of .83 (worry) and .69 (emotionality) for the W-EQ. Further alpha reliabilities in the .79 - .88 range for both scales of the W-EQ were reported by Morris and Fulmer (1976) and by Deffenbacher and Hazaleus (1985). Both the ITA and the WE-Q have a 5 point scale item response. Nonetheless, Osipow and Kreinbring (1971) employed a true/false version of the ITA and found a test-retest reliability of .68 for emotionality and .72 for worry over an eight week period. Deffenbacher (1980) could only obtain test retest reliabilities of .49 and .52 for emotionality and of .43 and .48 for worry, with the W-EQ.

Although more direct comparisons would have been desirable, on the basis of available evidence it appears that the ITA is psychometrically superior to the W-EQ. Moreover, the ITA items deal specifically with exams, and further it appears suitable to be administered even at some time after an examination has taken place, not necessarily straight after as the W-EQ demands.

Factors contributing to the choice of the ITA were its psychometric superiority, its "state test anxiety" nature, its reference to exams, and its more flexible administration criteria.

The ITA is further described in the next chapter (section 4.1.2).

It is expected that the ITA would have reflected test anxiety levels as predicted in the previous chapter. An analysis will be carried out to see if the two treatments which would presumably allay worry but not emotionality (PK and CATH) actually do decrease worry more than emotionality; however judging from the outcome of previous similar analyses (c.f., section 1.3) it is not expected that this step will yield significant results.

3.2 Test Anxiety Self-Efficacy

An important aspect of this investigation was the testing of the self-efficacy theory of behaviour change (Bandura, 1977). This theory states that any psychotherapy treatment alters the level and strength of self-efficacy, and further, that the "expectations of personal self-efficacy determine whether coping behaviour will be initiated, how much effort

will be expended, and how long it will be sustained in the ace of obstacles and aversive experiences "(p.191). In short he success of the treatment is a direct function of the legree to which it changes the client's self-efficacy feelings in the appropriate direction).

In testing this theory I was faced with two problems. Firstly, an appropriate measure of self-efficacy had to be constructed and, secondly, a criterion against which to assess the predictive validity of the self-efficacy measure had to be devised.

3.2.1 The first problem: Measuring self-efficacy

... was solved by constructing the Test Anxiety Self Efficacy Questionnaire (TASEQ) which has a scale for worry and one for emotionality and which is described in the next chapter (section 4.2).

The Measure of Academic Self-Efficacy (MASE) (Lalonde, 1979) was available but I chose not to use it for two reasons. One is that the MASE is a general measure of academic self-efficacy and therefore not sensitive enough for the purposes of this study. The other is that it appears to be measuring the performance aspect of self-efficacy, not the anxiety aspect. The discussion that follows will clarify what I mean.

A distinction in the conceptualization of self-efficacy is in order; to my knowledge, it has so far not been made: self

efficacy that one will not feel anxious in particular circumstances is different from self-efficacy that one will succeed on a given task.

The former draws on the ability to manage the anxiety spiral (or better, not to let such a spiral even begin), while the latter draws on one's aptitudes and skills required for succeeding on a given task. The latter is necessary for undertaking a task: performance on such task will be affected by both the former and the latter. The existence of the former presupposes the existence of the latter: one cannot be test anxious unless one takes a test or exam and taking a test or exam presupposes the belief in success to some degree.

A continuum between success and failure is assumed here and not a discrete distinction between the two. Self-efficacy managing anxiety (SEMA) and self-efficacy in succeeding (SESC) are two different although related constructs. Self-efficacy research so far has been predominantly of the SEMA kind (c.f. Bandura, 1982).

In view of this distinction it is not surprising that the correlations between the MASE and the TAS reported by Lalonde (1979) are low-between -.33 and -.39. To support my argument I should point out that those items reflecting the existence of test anxiety which were included in the provisional form of the MASE were subsequently eliminated from the final version failing to correlate highly enough with the total score.

The differentiation SEMA and SESC is obscured when we consider the treatment of phobias, which has been the main arena of self-efficacy research so far. This confusion is fuelled by the close association between the behaviour phobics and the degree of their phobia: the closer an aquaphobic can

get to the sea the less his/her phobia and vice versa. Following effective psychotherapy the aquaphobic will learn skills to manage his/her anxiety when approaching the sea and therefore his/her (SEMA) will be enhanced. The behaviour we would observe in this case would be walking towards the sea, but therapy has done nothing to influence the person's ability to walk - it was entirely within his/her repertoire. Approaching the sea becomes the criterion for treatment: the subject of treatment is enabling the person to control his or her anxiety (SEMA) while approaching the sea.

Suppose that the swimming association had established the "Far Prize" for the first person to swim to the "Far Island" and back within daylight hours. Aquaphobic Mark would not even consider swimming to the Far Island and back, let alone swimming it faster than anyone else to date. But again suppose that Mark's phobia could disappear for five minutes at the flip of the fingers - we could quickly measure his self-efficacy feelings that he could win the "Far Prize" (SESC) before those five minutes are up. Now let Mark go to a very good psychologist who effectively overcomes his aquaphobia (i.e., its related SEMA will reach the maximum level): Mark's SESC feelings in winning the Far Prize would be unchanged because they are based on his swimming skills, physical fitness, and motivation to achieve in this task and these have remained virtually unchanged by systematic desensitization or partition modeling or other techniques used by the very good pyschologist to help Mark overcome his aquaphobia.

The point I wish to make is that Mark's ability to approach the sea is a strong correlate of his phobia but with

weak causal link. Prior to Mark's visit to the very good psychologist he could have been induced to approach the sea by making important reinforcements contingent upon his approaching the shore, but in this case his aquaphobia would have probably remained unchanged.

A questionnaire designed to measure Mark's self efficacy pertaining to his ability to approach the sea would have items asking him how confident he was that he could, for instance, stand 10 meters from the sea. If his phobia was being overcome Mark would probably be confident that he could stand 10 meters from the sea (i.e., high self efficacy relating to approaching the sea). However, we could construct a self-efficacy questionnaire that made no reference to one's ability to approach the sea, but simply measured self-efficacy in managing one's own anxiety (SEMA) when a certain distance from the sea.

The problem with devising a traditional self-efficacy questionnaire for test anxiety is that the overt test behaviour of highly test anxious people is not obviously different from that of those who are low on test anxiety: they all go to the exam room and write, or tick boxes for about the same length of time. Moreover, we have seen in section 1.2.3 that the physiological arousal of high and low test anxious individuals does not seem to differ. What clearly differentiates the high test anxious individual from the low counterpart is the degree of anxiety experienced during test conditions.

In the light of the difficulty in discriminating between the behaviour of the high and low test anxious during test conditions, a questionnaire measuring self-efficacy in

anaging one's anxiety (SEMA) during exams appeared more ppropriate for this investigation. Hence the items forming he TASEQ deal mainly with one's confidence in managing anxiety during exams.

The reader with insight might have guessed. that the thrust of my argument is that the SEMA measurement of self efficacy is tantamount to an indirect way to measure a specific anxiety.

It is hoped that the distinction between SEMA and SESC discussed above will help to resolve Wine's (1980) paradoxical position of using self-efficacy as a unifying theoretical construct for test anxiety, yet acknowledging that betwen the two there is "some overlap but not identity" (p. 357) - a deduction made from Lalonde's (1979) reported low correlation, (-.36) between the MASE and the TAS.

As a note of caution I would like to add that although the reliability of MASE is good its validation has received only tentative support (c.f. Lalonde, 1979).

3.2.2 The second problem: Devising a criterion

As explained above, differentiating between low and high test anxiety behaviourally or physiologically has, to date, been very difficult if not impossible. So how can we validate the TASEQ and test Bandura (1977) self-efficacy theory in this area of research? We can ask subjects how effectively they managed their test anxiety during an actual exam and see how closely the effectiveness with which they managed such anxiety reflects their self-efficacy feelings that they would prior to

taking such exam. That is why the Post-Test Anxiety Self-Efficacy Questionnaire (P-TASEQ) was constructed.

Each item of the P-TASEQ is a past tense equivalent of TASEQ's. This strategy was employed so that this criterion would be as close as possible to what the self-efficacy questionnaire was meant to predict, thereby constituting a valid test for self-efficacy theory in this area of research.

The P-TASEQ is further described in the next chapter (section 4.2.2).

3.2.3 Rationale, hypotheses, and predictions behind the uses of TASEQ and P-TASEQ

The TASEQ was administered before and after treatment to establish whether SEMA feelings would parallel test anxiety decreases as measured by the AAT and the ITA. At the 4 week and twelve-month follow-up the TASEQ was again administered to see if SEMA feelings relating to taking tests and exams were maintained. The predictions with regard to the above hypotheses were that SEMA feelings as measured by the TASEQ will parallel test anxiety levels as measured by the AAT and the ITA. Moreover, it was predicted that increased SEMA feelings would be maintained at the two follow-ups. It was not expected that the worry and emotionality scales of the TASEQ would differentially vary but an analysis would be made to see if they had.

In addition to the four times outlined above, the TASEQ was administered prior to first and last final examinations. This time the P-TASEQ was also administered after these two examinations making it possible to test the self-efficacy theory. It was expected that TASEQ scores would reliably

predict P-TASEQ scores.

The first and last examinations were chosen to see if following treatment more experience at taking examinations enhanced feelings of SEMA and whether such practice improved subjects' estimation of experienced test anxiety, i.e. improved TASEQ prediction of P-TASEQ.

The TASEQ would be validated and it has expected that this endeavour would show it to be a valid and reliable measure of test anxiety self-efficacy.

3.3 Scholastic ability

A measure of scholastic ability, the Cooperative School and College Ability Tests-Series II (SCATT) (Educational Testing Service, 1967), serves three purposes.

One was to see whether performance on this ability test would improve following treatment and whether there was any difference between groups.

The second purpose was to find out whether aptitude acts as a moderator variable in the relationship between decreased test anxiety and improved academic performance.

A number of studies have found that high general anxiety facilitates performance of high aptitude subjects on cognitive-intellectual tasks (Denny, 1966; Gaudry & Spielberger, 1970, 1971; Katahn, 1966; Spielberger, 1966b). Kight and Sassenrath (1966) reported that high test anxiety facilitates performance on cognitive-intellectual tasks for students with high achievement motivation. Moreover, there is

considerable evidence (Spielberger, 1962, 1966b; Spielberger & Katzenmeyer, 1959; Spielberger & Weitz, 1964) which suggests a small but negative correlation between general anxiety and academic performance in the broad middle range of aptitude while no relationship was evident for subjects at the top or bottom of the aptitude distribution. Of greater relevance to this investigation is a study conducted by Paul and Ericksen (1964) as they employed a measure of test anxiety (the TAQ). As previously pointed out, Paul and Ericksen found that for students whose scholastic aptitude (measured by the SCATT) fell in the middle 70% of the SCATT distribution, test anxiety adversely affected their performance on a "real life" examination, but that no such effect was apparent for students at both extremes of aptitude.

In a sample of Grade 7 pupils Gaudry and Fitzgerald (1971) found that high test anxiety facilitates academic performance of the most able ones (i.e., those pupils in the top range of the academic ability distribution), while it debilitates that of the reminder; moreover, high test anxiety was found to be associated with the greatest performance deficit at the second highest of the 5 levels of ability.

Although the studies that have investigated the relationship between (test) anxiety and performance at various ranges of aptitude have not considered whether either the verbal ability or the mathematical ability range is the more crucial, this study did. Correlations between test anxiety and measures of academic and ability test performance will be obtained selecting: (a) those subjects whose score falls in the middle 70% of the total SCATT, (b) those subjects whose score falls in the middle 70% of the verbal scale of the

SCATT, (c) those subjects whose score falls in the middle 70% of the numerical scale of the SCATT. The same will be done with respect to changes in test anxiety measures and changes in academic and ability test measures.

I ought to point out that even though Cronbach and Furby (1970) argued that change scores are problematic for many research applications, extensive mathematical proofs have now shown that change scores are the best measure in controlled randomized trials (Berry, Bush, Olshen, Smallwood, & Kaplan, 1979; in McCordick et al., 1981).

If the correlations of interest appear to be stronger in any of the above three ranges of ability, then between groups comparisons testing for differential effectiveness in improving academic or ability test performance shall be crucial when those subjects falling in that range of ability are considered.

The third purpose of the SCATT was to see if either of its two scales moderate the effectivenenss of treatment in decreasing test anxiety.

3.3.1 Predictions and hypotheses

The predictions and inherent hypotheses pertaining to the SCATT are the following. (1) Subjects' performance on this test will improve significantly as a result of treatment. (2) There is a crucial ability range where the relationship between test anxiety and performance is stronger; it is in this range that the target (PK) group will show its

superiority more clearly. (3) The two scales of this ability measure may act as a moderator variable on the effectiveness of test anxiety treatment.

3.4 Study and Examination Skills

The principal reason why study and exam skills were measured in this study was to enable the evaluation of the study and examination skills (SES) program. The question advanced was: does this treatment improve subjects study skills?

In the test anxiety treatment literature it has been assumed that study skills treatments reduce test anxiety by improving subjects' study skills. In other words, it has been assumed that a study skills treatment improves group study skills - the evaluation of such treatment was often considered only in its effectiveness in reducing test anxiety, not in enhancing subjects study skills. Nonetheless, simple study skills training may have no effect on subjects' study skills (c.f. Jackson & Van Zoost, 1972).

In any case study skills programs in the test anxiety treatment literature have varied to such an extent that I believe it is important to evaluate their effectiveness in improving study skills as well as in decreasing test anxiety. In short, in our appraisal of the effectiveness of study skills programs in decreasing test anxiety and/or improving academic performance the ability of such programs in significantly enhancing subjects' study skills is an important prerequisite.

Another important consideration in employing a study skills measure was to see whether pretreatment study skills and examination skills acted as a moderator variable in the effectiveness of treatment in decreasing test anxiety.

Study and exam skills have been measured with the Inventory of Study Habits (ISH) (Jackson, Reid, & Croft, 1979). This measure appears to be highly reliable, and it has good construct and criterion related validity (c.f. Jackson et al., 1979). The ISH is described in the next chapter (section 4.3.2).

3.4.1 Predictions and hypotheses

The predictions and inherent hypotheses pertaining to study and exam skills are as follows.

(1) Subjects receiving SES treatments will show enhanced study and exam skills to a greater extent than other subjects. Implicit is the expectation that all subjects will improve their score on the ISH: greater familiarity with the inventory and subjects thinking and questioning their study habits following the completion of the ISH, the enhanced motivation to study due to attending the treatment sessions, and studying for final examinations (treatment ends a couple of weeks prior to finals) are all factors which are likely to improve subjects reported study skills at post-test and at the 4-week follow-up when the full ISH was administered.

(2) Study skills measures may act as moderator variables on the effectiveness of treatment in decreasing test anxiety.

3.5 Final Grades and Tests and Exam Performances

As one of the purposes of this study is to test the effectiveness of the target (PK) treatment in improving academic performance, two such measures were collected: (1) final grades obtained by subjects for the year before treatment, the year during which treatment took place, and the following year; (2) marks gained in tests and exams before and after (but not during) treatment.

Originally it was thought that the second (2) measure would have constituted a more stringent test for a test anxiety treatment, but it appears that highly test anxious students tend to do better in tests and exams held during the year than in final exams (Gaudry & Bradshaw, 1970) while the reverse is true for the low test anxious students. This finding is consistent with research reviewed in section 1.2.2.3 where we saw that stressful instructions produce deleterious effects on the performance of high test anxious individuals but facilitate performance of low test anxious individuals, and that under reassuring instructions the reverse happens. In short, as S. Sarason et al., (1960, p.136-158) hypothesised, the more "test-like" the situation, the more deleterious the effects of anxiety.

Moreover, Gaudry and Bradshaw's (1970) criterion for high or low test anxiety was a TASC score above or below the median. (The TASC was modified to be used with junior secondary school pupils, the subjects of their study.) It is likely that if they had selected a different criterion for high test anxiety (the top 13% of their test anxiety scale

distribution, for example) the effect they found could have been much stronger.

The relevance of Gaudry and Bradshaw's study to this treatment investigation lies in the fact that tests and exams before treatment were held during the year (i.e., they were internal exams) while the exams after treatment were all finals. Therefore, a pre-test versus post-test comparison on this index involved the behaviour of a variable which would normally show a decline. Moreover, it became apparent that some courses' internal exams are considerably easier than the final exam. Some subjects gained 100% and many gained marks around 80% in internal exams, but such high marks were unusual among final year exams. Consequently the validity of this measure in assessing a test anxiety treatment effectiveness in improving academic performance is highly questionable.

It was hypothesized that the PK group would come first when compared with the other groups on these two measures of academic performance. Moreover, it is hypothesized that the pattern of the PK group advantages over the other groups when academic ability ranges are considered will be as hypothesized in section 3.3.

3.6 Personality Variables

A number of personality variables have been measured in this study to see whether they predict the success of the various therapies in reducing test anxiety.

Although a long standing preoccupation (c.f. Meichenbaum, Gilmore, & Fedoravicious, 1971), assessing what subject characteristics are associated with the optional success of a therapy has received little attention in the psychotherapy literature. A handful of studies have identified several subject variables moderating the outcome of test anxiety treatments. Thus far general anxiety (McMillan & Osterhouse, 1972; Mitchell & Ingham, 1970), study skills (Vagg, 1978), social anxiety, emotionality and gender (Scrivner, 1974) have been found to moderate the course of test anxiety treatment.

Scholastic ability, study skills, and test taking skills, which have been discussed in the previous section, will be scrutinized as variables moderating the effectiveness of treatment. Originally the SCATT and the Cognitive-Somatic Anxiety Questionnaire (Schwartz, Davidson & Goleman, 1978) were the only measures that served as treatment moderating variables; consequently some of these measures were administered after treatment (at the 4-week follow-up). Specifically these were: the trait scale of the State-Trait Anxiety Inventory (A-Trait) (Spielberger, Gorsuch & Lushene, 1970); the Ray Achievement Motivation scale (Ray, 1979); and two measures of rigidity: The Einstellung Water Jug test (Rokeach, 1948) and an adaptation of the Ethnocentrism scale (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1969) to the white New Zealand population (all subjects in this study were

white New Zealanders).

The unfortunate consequence of administering these measures at the 4-week folllow-up is that they could have been modified by treatment. However there is no reason to believe that those personality characteristics measured by the above four scales would be altered by any of the four treatment procedures with one exception: trait anxiety. Even though test anxiety treatment studies employing systematic rational restructuring (Goldfried, Lineham & Smith, 1978) and cognitive attentional training on its own or in addiion to relaxation (Wine, 1970; in Denney, 1980) have failed to decrease trait anxiety (on the same measure employed here) it is possible that one or more of the treatment procedures employed in this study did so. Notwithstanding the above concern, a more or less uniform change in subjects' trait anxiety following treatment would not affect the result of the analysis of covariance which will be used to establish whether trait anxiety acts as a moderator variable between the effects of treatment on test anxiety and performance improvement.

An indication of whether trait anxiety can be confidently employed in the intended manner can be obtained by performing a correlation between the cognitive scale of the CSAQ (CSAQ.C) and the A-Trait; as Schwartz et al. (1978) found the two to yield a very significant ($p < .001$) and high correlation ($r = .67$) (the corresponding correlation for the somatic scale (CSAQ.C) was .40). This correlation will give an indication of the stability of subjects ranking between the CSAQ.C and the A-Trait. Given that these two scales correlate highly, should the above correlation be about as high there are two possibilities pertaining to A-Trait scores: either (i) they

did not change from pre-treatment levels or (ii) the change was fairly uniform. In either case we can be confident that the moderator variable analysis discussed above for A-Trait is valid.

Correlations between personality variables and the test anxiety measures (at the four different measurements) and intercorrelations between personality variables shall be examined to see if any pattern is apparent.

3.6.1 The Cognitive-Somatic Anxiety Questionnaire

Schwartz and associates (1978) reviewed a body of research suggesting that anxiety manifestations follow certain specific patterns rather than being an overall response to perceived threat. They developed a questionnaire that measures anxiety in its cognitive and somatic manifestations: The CSAQ. Schwartz et al. argued that the subdivision of anxiety between cognitive and somatic "is the most basic of all splits" and that further differentiations are possible. For example somatic may be further subdivided into skeletal and autonomic, and cognitive into right versus left hemisphere mediated.

These researchers also argued that one ought to tailor anxiety treatment to match the individuals' reactions to anxiety. The long range aim for the employment of the CSAQ in this study was just that: to find out what sorts of treatment best suit a client of known typical reactions to anxiety. In a shorter perspective the rationale for the inclusion of the CSAQ was to find out whether subjects high on either the CSAQ.C or the CSAQ.S would benefit more from any of the

treatment procedures employed in this investigation.

It was predicted that the subjects high on CSAQ.C would benefit from treatment more than subjects low on this measure.

The CSAQ is described in the next chapter (section 4.4.1)

3.6.2 The Trait scale of the State-Trait Anxiety Inventory

McMillan and Osterhouse (1972) found that desensitization significantly reduced self-reported test anxiety, but that only for those subjects low on general anxiety (as measured by the MAS) were improvements in academic performance acheived.

Mitchell and Ingham (1970) also employed desensitization as a procedure to reduce test anxiety, and controlled (statistically) for general anxiety as measured by the IPAT general anxiety scale. At post-test it was found that this treatment seemed to be more effective at decreasing test anxiety for the high general anxiety subjects, but that they lost such advantage at the 14-week follow-up. It is important to note that subjects in this study were academic failures. Mitchell and Ingham did not measure the effects of treatment on performance but they noted that low general anxiety subjects increased their score on the AAT+ more than the high general anxiety subjects. This effect was not significant ($p < .12$) but it was very strong. Bearing in mind the moderate positive correlation between the AAT+ and academic performance (e.g., Alpert and Haber, 1960), it is possible that if these researchers had measured the effects of treatment on academic performance, they would have obtained the same result as

McMillan and Osterhouse, (1972).

Although different types of treatment were used in this study it was predicted that subjects low on the A-Trait would achieve greater performance improvements as a result of test anxiety decrements than subjects with high A-Trait levels. It is also predicted that subjects low on A-Trait would report greater test anxiety reductions than the high counterparts. The rationale for this prediction was that irrelevant and negative thinking in testing situations (i.e., test anxiety) is more deeply ingrained in the cognitive structures (Meichenbaum & Butler, 1980) of high A-Trait than low A-Trait individuals and therefore less easily eradicated.

The State-Trait Anxiety Inventory is described in the next chapter (section 4.4.2).

Should there be evidence that trait anxiety was affected by treatment, the CSAQ scales would be employed in place of the A-Trait scale.

3.6.3 The short form of the Ray Achievement Motivation Scale

It was reasoned that higher levels of achievement motivation would facilitate the treatment process. Highly motivated subjects would presumably complete homework assignments, be attentive to the therapist's guidance, and follow his suggestions with greater efforts, thereby deriving greater benefits from treatments than the less motivated ones.

A short and quick-to-complete form of the Ray Achievement Motivation (RAM) scale (Ray, 1979) was sought as subjects' already had to complete a long list of questionnaires.

It was predicted that subjects with high RAM scores would benefit more from treatment than subjects with lower RAM scores.

The RAM is described in the next chapter (section 4.5).

3.6.4 The two rigidity measures

Because the treatment programs employed in this study were largely or entirely cognitive and because the success of therapy demanded that subjects change their thought patterns and cognitions and/or their academic habits, I reasoned that high levels of mental rigidity would hinder normal therapeutic progress. Subjects who were excessively fixed (rigid) in their thinking patterns would find it difficult to change their self-talk and academic habits and so derive less benefit from a particular treatment program.

As I have outlined in the next chapter (section 4.6), much controversy and little conclusion surrounds the rigidity literature. Most problematic of all is its measurement. Different measures appearing in this literature, all presumably of rigidity, failed to correlate significantly (Applezwig, 1954; Goodstein, 1953). This posed two fundamental questions: (1) Does rigidity exist? and (2) whatever it is that these scales measure - which was the best index? I chose to use the Einstellung Water Jug (EWJ) test (Rokeach, 1948) and the Ethnocentrism scale (Adorno, et al., 1969) because they were found to yield virtually a zero correlation (Applezwig, 1954; Goodstein, 1953) and because

they are very dissimilar. One is a type of arithmetic test and the other a questionnaire measuring prejudice towards the ethnically unlike. Because they were measuring entirely different things (as suggested by their failure to correlate at all) it was more likely that either of the two would measure rigidity.

It is predicted that rigidity will hinder the therapeutic effect of the treatment.

The EWJ test and the Ethnocentrism scale employed in this study are described in the next chapter (section 4.6)

3.7 Social Validity

I have pointed out in the previous chapter (section 2.3) that the social validity (Wolf, 1978) of test anxiety treatments reported in the literature has been disturbingly neglected. With few exceptions (e.g., Holroyd, 1976), treatments were evaluated solely on the basis of test anxiety questionnaires data as well as performance measure in about half the cases.

Briefly, Wolf (1978) identified three levels of social validity: the social significance of treatment goals, the social appropriateness of the procedures, and the social significance of their effects. Having assumed that the goals of this treatment study (reducing subjects' test anxiety and improving their grades) were socially desirable, I was interested in learning whether the treatments were well received by those involved and whether they felt the treatments were beneficial and to what extent. In this sense

the two social validity questionnaires pertained to the measurement of something that was akin to participant reactions, a category of training evaluation criteria (Kirkpatrick, 1967; Kane, 1976).

The social validity questionnaires, namely The First Impression of Treatment (FIT) questionnaire administered at the end of the first treatment session, and the Participant Reaction questionnaire (PRQ) administered at post-test and at the 4-week follow up, serve various important purposes. The first was to see whether the four treatments were well received and whether they were perceived to be of equal therapeutic potential. Moreover, it was going to be investigated whether treatment expectancy acted as a significant moderator variable on the effectiveness of treatment at decreasing test anxiety. It has been assumed that subjects' positive expectations of improvement influence the effectiveness of therapy but no one to my knowledge has investigated this assumption empirically.

The second purpose of the social validity questionnaires was to validate the effectiveness of test anxiety measures. This was done by checking if the degree of test anxiety reduction indicated by the scales employed would match the degree of test anxiety reduction reported by the subjects when asked directly. The third important purpose was to assess the subjects' perception of the therapist in the various treatment groups. This was important in ascertaining the hypotheses with respect to the four treatment groups. He could have wittingly or unwittingly tried harder to alleviate the test anxiety of subjects in the PK group, in which case it may not have been

the actual treatment that made the difference but, the therapist's degree of effort and enthusiasm that did so. In short, one of the reasons for these social validity questionnaires was to see whether non-specific treatment effects (Kazdin & Wilcoxon, 1976) were present in varying degrees among the four groups. Nonetheless, considering that participants were intelligent people it was more likely that subjects in the PK groups would find that the therapist would have "understood their particular problem" more accurately and would find their program more credible than subjects in the other groups. As Wilkins (1985) has argued, credibility is an inherent aspect of a therapy.

The PRQ was readministered at the 4-week follow-up to see if subjects' opinion of treatment had changed after the 4 weeks during which they sat their final exams and had plenty of opportunity to put into practice what they had learned during the treatment.

It is predicted that subjects perceptions of the four treatments as measured by the FIT and the PRQ will not be significantly different; furthermore, that subjects' opinion of their treatment will not change significantly after 4 weeks.

The FIT and the PRQ questionnaires are described in the next chapter (section 4.7).

3.8 Follow-up Data Collection

Briefly stated, the rationale for the collection of data at 4 weeks and at 12 months after treatment was to

see whether the effects of the four treatment procedures, evident at post-test, would be maintained over time.

It is predicted that treatment effects will be maintained at the two follow-ups.

CHAPTER IV

PARAMETERS, EXPERIMENTAL VARIABLES, AND MEASUREMENT.

Measurement is so crucial an aspect of scientific research that it can be argued that the quality of the latter depends on the validity and reliability of the measurement instruments employed.

Variables from seven parameters were examined in this treatment investigation: (1) test anxiety, (2) self-efficacy, (3) scholastic variables, (4) general anxiety, (5) achievement motivation, (6) rigidity, and (7) social validity.

A feature common to all but two of the instruments employed is their susceptibility to faking, either good or bad. The Cooperative School and College Achievement Test - Series II and the Einstellung Water Jug test can only be faked bad.

4.1 Test Anxiety

Two measures of test anxiety are employed in this investigation: the Achievement Anxiety Test (Alpert & Haber, 1960) and the Inventory of Test Anxiety (Osterhouse, 1972, 1976).

4.1.1 The Achievement Anxiety Test

The Achievement Anxiety Test (AAT) was developed and first

validated by Alpert and Haber (1960). Dember, Nairne, and Miller (1962), Kirkland and Hollandsworth (1979), McKordick, Kaplan, Smith, and Finn (1981) and Milholland (1964) have provided further evidence as to its validity. The AAT is becoming the single most widely used test anxiety instrument. Its most distinctive feature remains the differentiation between debilitative and facilitative anxiety, rejected in its two scales, AAT- and AAT+ respectively.

I am surprised that no controversy has arisen over what these two scales actually measure. It has usually been assumed that facilitative anxiety (AAT+) facilitates performance while debilitative anxiety (AAT-) debilitates it. However the reasoning behind this assumption is inexorably circular as its authors chose items which correlated negatively with measures of academic performance and verbal aptitude for the construction of the AAT- but positively for the construction of the AAT+.

Another criterion adopted by Alpert and Haber in their construction of the AAT was that items from each scale correlated least with items from the other scale. In spite of these efforts the AAT- and the AAT+ correlated significantly in all samples employed by their authors, yielding an average correlation of -.37. Interestingly enough, studies that followed (e.g., Dember, Nairne, & Miller, 1962; Hendel, 1980; Huck & Jacko, 1974; Kirkland & Hollandsworth, 1979; Pervin, 1967) consistently obtained higher such correlation (ranging between -.42 and -.66), suggesting that statistical regression (Cook & Campbell, 1979) was operating.

Weiner and Samuel (1975) provided evidence which suggests that whether anxiety is facilitative or debilitative depends

on the labeling of one's arousal. It is clear that the distinction between facilitative and debilitative anxiety is in practice not as clear cut as the AAT purports. Moreover, there is evidence which suggests that low test anxious individuals label their arousal as facilitative and high test anxious individuals label it as debilitative (Hollandsworth, Glazeski, Kirkland, Jones, & Van Norman. 1979), while, as we found in section 1.2.3, their degree of arousal does not seem to differ significantly (C. Brown & Gelder, 1938; Deffenbacher & Hazaleus, 1985; Hollandsworth et al., 1979; Holroyd & Appel, 1980; Holroyd, Westbrook, Wolf, & Badhorn, 1978).

There are nine items in the AAT- scale and ten in the AAT+, which with nine "buffer items" add up to a 28 - item instrument. The subjects answer each item on a 5 - point scale, indicating the degree to which it applies to them.

As outlined above the 19 items of the AAT scales were content validated as well as being examined for criterion-related validity. Moreover their reliability appears to be good: Alpert and Haber report test-retest reliabilities for a 10-week interval of .83 (AAT+) and .87 (AAT-) and of .75 (AAT+) and .76 (AAT-) for an 8-month period. Huck and Jack (1974) employed three different versions of the AAT (see below) and obtained alpha reliability coefficients ranging between .65 and .73 for the AAT+, .76 and .86 for the AAT-, and .83 and .88 for the AAT+ minus the AAT- (AAT+-) scale.

Alpert and Haber found that both scales used together are better predictors of academic performance than either scale alone. This result was corroborated by several other studies that followed (Carrier & Jewell, 1966: Huck an Jacko, 1974; Kirkland & Hollandsworth, 1979).

Huck and Jacko (1974) and Kirkland and Hollandsworth (1979), two rarely cited studies (a fact which the writer believes is very unfortunate), provide valuable information on this measure of test anxiety. The authors of both these studies argue that the difference between the two scales is a better measure of test anxiety than either of the two alone. Furthermore, evidence in Kirkland and Hollandsworth's study suggests that although highly correlated, the AAT is superior to the TAS.

Huck and Jacko's findings are very pertinent to this study. These two authors observed that researchers were sometimes using not the original AAT but altered forms of it. They noted that this fact was not pointed out in the relevant journal articles, nor was there any description of the alterations made. What Huck and Jacko did was to compare the original form of the AAT with the two altered ones. One of the alterations had been made by Smouse and Munz (1969), the other by Walsh (1968, 1969) and Walsh, Engbretson and O'Brien (1968). Each of the 28 items of the original AAT had a five-response alternative listed vertically, in a multiple choice format. For example the sixteenth item appears as follows:

I look forward to exams

(a) Never
(b) Hardly ever
(c) Sometimes
(d) Usually
(e) Always

Throughout the original AAT the responses were tailored to each statement (including the buffer items) and thus the phraseology of the responses varied across the 28-item

inventory.

The two altered versions investigated by Huck and Jacko varied in that neither included the nine buffer items and, further, their response formats were different from each other and from the original. Smouse and Munz used a 5-point horizontal continuum with its endpoints verbally defined by the same words or phrases that were used for the first and last responses in the original AAT. Thus Item 16 in this version appeared as follows:

I look forward to exams

1 2 3 4 5
Never Always

Walsh et al., similarly, did not include the nine buffer items and used a Likert-type response format for each item with the phraseology of the five response options constant across all items. Thus item 16 in this version appeared as follows:

I look forward to exams R S F G A

The directions explained that R = rarely, S = sometimes, F = frequently, G = generally, A = always.

Huck and Jacko found no substantial difference between the original AAT and the variations used by Walsh et al. However, scores on the Smouse and Munz version were significantly and substantially lower, yielding means of 19.288 and 23.173 as opposed to the original scores of 24.933 and 29.356 for the AAT+ and AAT- respectively (i.e., about six points lower than the original version. Standard deviations reported by Alpert

and Haber (1960), Huck and Jacko (1974) and Kirkland and Hollandsworth (1979) are comparable for both scales, at about 5 or 6. The reliability estimates for the three AATs were also very comparable. Now, given that the mean for the Smouse and Munz version one is about 5 or 6 points lower and that the standard deviation is of very similar magnitude to that of the original AAT, a given raw score on this modified version will correspond to a much higher percentile rank.

The pertinence of Huck and Jacko findings to this investigation lies in the fact that I employed a version of the AAT which is very similar to the one employed by Smouse and Munz. (The exact form can be found in Appendix A.)

It appears that Holroyd (1976) also did not use the original AAT but the Smouse and Munz version or one that yields very similar results. He reported that "previous administrations of the AAT to students taking introductory psychology courses indicated that a cut-off score of 32 included approximately the upper 13% of the distribution of scores on this measure "(i.e., the AAT-) (p.992). Data available on the original AAT would place a raw score of 32 in the debilitative scale only about half a standard deviation above the mean.

Snyder and Deffenbacher (1977) have probably used a similar version as they found that a cut off score of 30 included the upper 15% of scores in a sample of 350 college students and used it as a criterion for selecting subjects in their test anxiety treatment investigation.

Following Snyder and Deffenbacher's (1977) findings, Harris and Johnson (1980, 1983) and Wise and Haynes (1983) used this very subject selection criterion for their treatment

investigations. Huck and Jacko report a mean of 29.356 for the AAT- when using the original version: a cut off score of 30 would include approximately the upper 50% (not 15%). For the Smouse and Munz version the corresponding mean was 23.173 with a standard deviation of 6.606: a cut-off score of 30 on this version would indeed include approximately the upper 15%.

I think the reason for the apparent popularity of the Smouse and Munz version lies in the fact that it can be derived from the original article (Alpert & Haber, 1960) thus avoiding the extra effort required to obtain the full scale from its authors. Furthermore, I believe that the characteristics of this version which differ from the original are due to the following factors: (a) Alpert and Haber did not publish the nine buffer items of the AAT but only the ten forming the AAT+ and the nine forming the AAT-; they also indicated the corresponding item number in the original AAT for each of the 19 items published. (b) They gave the first and last response but not the middle three for each of the 19 items. (c) Such first and last responses were separated by a dash in between which is likely to give the impression (it happened to me) that those responses were the endpoints of a continuum. (d) The responses were to give an indication of the frequency with which the items contained were experienced (e.g., always - never), so their relative position also indicated whether the item had to be reverse scored or not. (e) In their description of the AAT, the authors stated: "The Ss answer each item on a FIVE-POINT-SCALE indicating the degree to which the item applies to them" (Alpert & Haber, 1960, p.213, capitals mine) - which supports one's likely but erroneous impression that the response format in the original

AAT is the more conventional scale with five equidistant points on a continuum. And (f) The likely belief that the buffer items are not very important which would probably be facilitated by the extra effort needed to obtain them.

My efforts to obtain the original AAT before subject selection for this study was due to begin were unsuccessful, and therefore I resorted to the use of the format reported in appendix A.

This confusion pertaining to such a popular test anxiety instrument has rather unfortunate consequences. Harris and Johnson (1980, 1983) used a cut-off score of 30 in the AAT- as a selection criterion in their treatment studies. Now, in my efforts to obtain the original AAT, I wrote to these authors for help and Gina Harris sent me a form of the AAT. Correspondence being slowed down by their holiday period, I received it too late to use it for selecting my subjects.

However, this version again varied from the original. Although the buffer items were included the response format was a fixed 5-point scale: (1) almost never, (2) rarely, (3) occasionally, (4) often, (5) almost always. The layout of this form of the AAT is similar to the popular State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970). Harris and Johnson (1980, 1983) based the choice of their cut-off score on the results obtained by Snyder and Deffenbacher (1977), but there is no indication as to which form had been used in the first instance.

This could lead to serious shortcomings. Suppose Snyder and Deffenbacher used Smouse and Munz's version (which seems reasonable in view of their results compared with those Huck and Jacko obtained) and that Harris and Johnson's version

yielded results similar to the original AAT (which does not seem unreasonable in view of the fact that the buffer items were included), we can see that Harris and Johnson's cut-off score could easily include the top 50% of this distribution and not the top 15% as intended. A treatment which is effective in decreasing test anxiety in moderately test anxious individuals could work differently or not be as effective with highly test anxious individuals.

As well as the AAT, Harris and Johnson (1983) used the STABS and the TAS to measure test anxiety changes. Their reported group means on the STABS are 140, 133, 161, 135, and 144, which taking into account not normative data published in Suinn (1969), reflect only average or moderately high levels of test anxiety. This gives support to my earlier reservations as to the appropriateness of Harris and Johnson's (1980, 1983) cut-off score as a criterion for subject selection. Harris and Johnson's (1980) group means on the STABS were 143.8, 136.3, 127.0, 142.8, and 156.0, which again reflect only average or moderately high levels of test anxiety. The TAS group means (23.0, 26.5, 21.1, 26.0, 26.0 for the 1980 study and 23.8, 23.9, 26.3, 24.0, 22.5, for the 1983 study) are difficult to interpret, although particularly for some of the treatment groups they appear to show only moderate and not high levels of test anxiety. Sarason (1978) reported means (and standard deviations) on the TAS of 19.74 (6.73) for females (n = 237) and of 16.72 (7.12) for males (n = 283). We also need to consider that, although Harris and Johnson (1980, 1983) did not report their group composition by sex, unless systematically recruited the great majority of test anxiety treatment subjects are female.

The statistical properties of the AAT employed in this study appear in all respects to be the same as in Smouse and Munz's (1969) version. According to Huck and Jacko they are the following; means 19.288 (AAT+), 23.173 (AAT-), -3.982 (AAT+-); standard deviation: 6.029 (AAT+), 6.606(AAT-), 10.627 (AAT+-); alpha reliabilities: .76 (AAT+), .76 (AAT-), .83 (AAT+-); AAT+ and AAT- correlation is -.42. The actual index derived from the AAT which has been employed in the statistical analysis is the following: AAT- minus AAT+ (AATD-F), this rather than AAT+ minus AAT- retains the qualities of test anxiety measures, that is, a lower score indicates less test anxiety.

The strength of the AAT has inspired test anxiety treatment researchers McKordick, Kaplan, Smith, and Finn (1981) to write the following in their discussion,

> The strongest results were for the self-report AAT scales. For the population we have been working with, the AAT HAS THE MOST IMPRESSIVE RECORD OF CONVERGENT AND DISCRIMINANT VALIDITY AMONG THE MANY MEASURES WE HAVE USED. The facilitative component correlates negatively with self-report anxiety measures, (range -.44 to -.63) and the debilitative component correlates positively (range .38 to .57). Further correlations between the AAT facilitating and debilitating scales with GPA were .22 and -.25 respectively. Although these associations are moderate in magnitude, other self-report measures do not correlate as highly. Indeed the AAT predicts GPA at our institution nearly as well as the Scholastic Aptitude Test (SAT). The AAT facilitating predicted performance on a particular mid-term exam (r = .32) nearly as well as GPA (r = .44) (p. 177, capitals mine).

And may I add: more power to Alpert and Haber and their AAT!

4.1.2 The Inventory of Test Anxiety

The Inventory of Test Anxiety (ITA) was developed by Osterhouse (1972, 1976); it has an 8-item scale for worry and one of equivalent length for emotionality, making a total of 16 items. Subjects answer each item on a 5-point scale; the possible range of scores on each scale is 8 - 40 while for the total ITA the range is 16 - 80. A copy of the ITA used in this study can be found in Appendix B.

The ITA has a strong affinity with the Worry Emotionality Questionnaire (WEQ) (Liebert and Morris, 1967). Osterhouse included all ten W-EQ items in the original list of 21 items from which the ITA was developed. Some were derived from different inventories, and still others were composed by Osterhouse himself. This inventory was content validated, its items were selected on the basis of 100% interjudge agreement regarding the dimensions involved, namely worry or emotionality.

Five of the original 21 items were eliminated. The remaining items were worded to refer only to classroom examinations just taken or taken in the recent past (which makes the ITA very suitable for this investigation). The person completing the questionnaire indicates to what degree the feeling, state, or condition described by each item was experienced during the recently taken examination on the 5-point scale.

The two scales of the ITA appear to have reasonably good reliabilities. Osterhouse (1972) obtained a .92 split-half

reliability (adjusted for length). Osipow and Kreinbring (1971) employed a true/false version of this questionnaire and found the following test-retest reliabilities for worry and emotionality respectively for five different groups: .65 and .63, .84 and .47, .73 and .62, .79 and .72, .72 and .68, for periods of approximately 4, 5, 6, 7, and 8 weeks respectively. On the whole these reliabilities may not appear very great, however we must consider that the ITA asks you to indicate how you actually felt during your last examination. Subjects are therefore likely to remember less accurately how they felt during a particular exam several weeks later. Moreover, we must consider that some students might have sat another exam during the retest interval and because the ITA is a measure of "state test anxiety" (c.f. section 3.1.3) this is likely to fluctuate from exam to exam.

4.2 Self-Efficacy

There are two variables that fall under this parameter: test anxiety self-efficacy and experienced test anxiety self-efficacy. The former is measured by The Test Anxiety Self-Efficacy questionnaire and the latter by the Post Test Anxiety Self Efficacy questionnaire.

4.2.1 The Test Anxiety Self-Efficacy Questionnaire

The rationale behind the construction of the Test Anxiety Self-Efficacy Questionnaire (TASEQ) has been discussed in the

previous chapter (section 3.2). The TASEQ was devised to measure self-efficacy feelings of managing one's anxiety (SEMA) during examination conditions (see section 3.2.1 for the distinction between SEMA and SESC in self-efficacy conceptualization). Accordingly TASEQ items do not deal with one's confidence in a task but deal with one's confidence that certain feelings, thoughts, and fears characteristic of test anxiety will not be bothersome or that they will not occur. As previously pointed out the nature of test anxiety and its impact on behaviour are not as conspicuous as when considering phobias. Behaviourally anchored items would not therefore be appropriate for the construction of a self-efficacy measure of test anxiety.

The format of the TASEQ is similar to self-efficacy questionnaires supplied to the author by Albert Bandura. The TASEQ consists of 21 items either derived from test anxiety scales (AAT, TAS, ITA, TAI) or devised by myself. Fourteen of them (1-13,21) form the worry scale and 7 (14-20) the emotionality scale.

Subjects rate their "level of confidence in" each particular item. The "confidence level" is rated on a 10-point scale ranging from 10 (quite uncertain) to 100 (certain). Items left blank imply that the subject cannot "do" them and yield a zero. In effect, for scoring purposes, the "confidence level" score is an 11-point one ranging from zero to 100. The total score ranges from zero to 2100, the worry scale from zero to 1400, and the emotionality scale from zero to 700. The greater proportion of worry items reflects its greater importance over emotionality in test anxiety. A higher score on the TASEQ indicates higher levels of SEMA

feelings relating to examinations and therefore to test anxiety.

In order to familiarize subjects with the rating format of the TASEQ a "practice rating" sheet is included. This sheet contains 14 items asking for one's confidence in lifting various weights of increasing magnitude. The "practice sheet" was placed before those with the items of interest but it was removed after the second administration of the TASEQ as subjects displayed considerable familiarity with it.

A copy of this questionnaire can be found in Appendix C.

4.2.2 The Post Test Anxiety Self Efficacy Questionnaire.

As pointed out in the previous chapter (section 3.2) the need for the P-TASEQ arose from the nature of self-efficacy theory testing; the degree of confidence that one could "do" a particular item has to be matched by actually "doing" what that particular item entails.

Also pointed out in section 3.2 was the difficulty (if not impossibility) of having overt behaviours as criteria for testing the predictability of the TASEQ. Therefore a self-report measure dealing with experienced anxiety during an exam and composed of the same TASEQ items but in the past tense was devised: the Post Test Anxiety Self-Efficacy Questionnaire (P-TASEQ).

The first item of the TASEQ, "I would feel I will do alright on this exam", in the P-TASEQ appears as "I felt I was going to do alright on this exam."

Subjects were required to rate the degree of accuracy of each item as it applied to them on a 10-point "accuracy scale"

ranging from 10 (quite inaccurate) to 100 (accurate). Those items left blank were counted as 100. It was reasoned that if a particular thought, feeling, or fear typical of test anxious students did not happen then it could not contribute to experienced test anxiety. A corollary to the above is that the accuracy rating of each item in the P-TASEQ indicates its relative contribution to the level of experienced test anxiety; a low accuracy score means that greater test anxiety was experienced on that given exam or test, and vice versa for a high accuracy score.

Like the TASEQ, the P-TASEQ contained 21 items. Of these 14 (1-13, 21) form the worry scale and 7 (14-20) form the emotionality scale. The possible range of scores was 140-1400 for the worry scale 70-700 for the emotionality scale, and 210-2100 for the total P-TASEQ.

A copy of the P-TASEQ can be found in Appendix D.

4.3 Scholastic Variables

Eight scholastic variables were considered in this investigation. Three of them were derived from the Cooperative School and College Ability Tests - Series II (SCATT) (Educational Testing Service, 1967): the verbal ability subtest, the numerical ability subtest, and their sum. Three more scholastic variables were derived from the Inventory of Study Habits (Jackson, Reid, & Croft, 1979). The first, which I have called Studying Habits, is the sum of the first 5 scales of this inventory: the Place of Study, the Study Times, the Organization for Study, the Textbook Reading Skills, and the Taking Notes. The remaining two scales,

Studying for Examination and Examination Technique, constitute the other two variables derived from the Jackson et al. (1979) inventory. The other two variables falling under this parameter are final grades, and tests and exam marks.

4.3.1 Cooperative School and College Ability Tests —Series II

The SCATT is a measure of scholastic ability comprising two subtests of 50 items each: the verbal ability subtest (V-SCATT) and the numerical ability subtest (N-SCATT). The series provides tests at four levels of difficulty with either two or three equivalent forms at each level. In this study the highest level tests (4) were employed. Form B was administered before treatment began and form A at the 12-month follow up.

All SCATT items have a 4-option answer with the exception of the lowest level test (1) where the SCATT-N items have a 3-option answer. The items in the verbal subtest consist of a target pair of words that "go together in a certain way" and four other pairs of words listed under it, the correct answer is the pair of words whose relationship is the same as the target's. The mathematical subtest items consist of two arithmetic or algebraic expressions for which subjects have to decide on one of the four options: (a) the first is greater than the second, (b) vice versa, (c) the two are equal, (d) there isn't enough information to decide. The possible range of scores for the total SCATT is 0-100 (0-50 for each subtest).

Reliability coefficients reported in the manual are .90 or greater for the mathematical subtest, .87 or greater for the verbal subtest, and .94 or greater for the total SCATT. Also reported in the manual are moderate coefficients for both predictive and constructive validity.

Reviews of this test tend to be favourable (e.g., Fox, 1978; Rosenbach 1978).

4.3.2 The Inventory of Study Habits

As pointed out above three variables are derived from the seven scales of the Inventory of Study Habits (ISH). Each of the seven scales comprises 25 items giving a total of 175. Subjects are asked to rate the frequency with which they engage in the behaviours described in each of the inventory's items on a 5 point scale which is the same for every item: never or almost never (N), about 1/4 of the time (1), about 1/2 the time (2), about 3/4 of the time (3), always or almost always (A). Eighty six of the 175 items are reverse scored, for these N = 2, 1 = 1, and 2,3, A = 0; for the remaining 89 A = 2, 3 = 1, and 2,1, N = 0. The possible range of scores is 0 - 50 for each of the scales and 0 - 350 for the total ISH.

The name of each scale gives a clear indication of what aspect of academic behaviour it is intended to measure: (1) Place of Study, (2) Study Times, (3) Organization for Study, (4) Textbook Reading Skills, (5) Taking Notes (the sum of which form the Studying Habits (SH) variable), and Studying for Examinations (SE) and Examination Technique (ET).

Reliability coefficients reported in the manual are quite

high (around .80) and intercorrelations among the seven scales are usually moderate with some of them being low. The manual also shows data supporting the validity of the seven scales; their zero-order correlations with English and three best subjects (E3B) were all moderately low and significant, but were low or not significant with the Otis Test of Mental Ability (OTIS). However multiple correlations of each ISH scale and the OTIS with E3B yielded high and significant coefficients (.70 to .75, all $p < .01$). This indicates that each of the ISH scales accounts for considerable academic performance variance not accounted for by IQ.

Unfortunately the manual offers no such data for the total ISH score. One would expect the total ISH to account for considerably more academic performance variance than either of its scales alone.

The ISH was developed for high school pupils and some of its items clearly referred to "school", but it was pointed out to the subjects that they should not be "offended" by some ISH items low face validity. It is not expected that this characteristic has affected SH, SE, or ET scores.

4.3.3 Final grades and tests and exam marks

Final grades gained by subjects the year (1983) when treatment took place, as well as for the year before and the year after, were collected. As these are typically recorded as letter grades they were transformed into a mark out of 100 so that they could be easily compared statistically. The

following conversions were made: A+ = 86, A = 77, A- = 74, B =71, B- = 60, C+ = 57, C = 52, C- = 49, D+ = 46, D = 42, and E = 30. Final grades from all courses taken each particular year were collected and averaged out to give a Final Grade Average (FGA) each year for each subject.

Unfortunately it was the first year at university for many subjects who did not have a 1982 FGA. Employing an FGA based on their performance in their final year at high school was considered for these subjects but not carried out as such a measure would not be equivalent to an FGA based on grades gained at university.

A similar strategy was employed with respect to marks gained by subjects on tests and exams held during the two terms prior to treatment and in final exams respectively. Letter grades and marks were all transformed into a mark out of 100 to allow the appropriate statistical comparisons. Letter grades were transformed using the same scale as for the final grades. All marks gained by a subject on tests held during the 1983 first and second term were collated and averaged out to give one Tests Exams Average (TEA). The same proceedure was followed for marks gained in final exams that year.

Data relevant to TEA had to be obtained from records held by each academic department concerned as it was not available elsewhere. Although some departments were very cooperative, this task proved to be a long and laborious one by concerned persons being out, poor record keeping, and suspiciousness about releasing marks. These were all unforeseen factors which made this task seem arduous.

Notwithstanding these difficulties, marks from tests and

exams for all courses undertaken by each subject were obtained with three exceptions. First, records for one course were not clearly identifiable with the subject concerned; second, one department was so concerned about the release of this information that the author decided not to obtain that piece of data; third a lecturer who held the relevant records was away on study leave at the time of data collection. Nonetheless it is unlikely that these would affect those three students TEAs to any significant extent given that (1) in each case only one subject was concerned (2) the remaining relevant data for each subject concerned had been obtained, and (3) marks gained by a subject in a given exam or test tend to be about as high as those gained by him/her in the rest of the exams.

Tests and exams held during the third term, when the four treatment programs took place, were not considered.

Standardized marks or students' percentiles for each exam or test would have been better measures to use in this study, but unfortunately only two departments calculated either standardized scores or students ranking on a test relative to the rest of the class, so raw marks had to be employed.

It was pointed out in the previous chapter (section 3.5) that there were considerable inconsistencies relating to this measure: some departments appeared to set considerably easier questions for internal exams.

Because some courses don't have internal exams and rely on essays, lab reports, and the final exam for assessment, a situation could arise where a subject would have only one or two internal exams. A high mark coming from one of the "easy" internal exams would be a source of limited error when

averaged out (to obtain TEA), with other not "easy" internal exams. However if only one or two internal exams were sat by a subject during that year, and if these came from a department that set "easy" internal exams, then the subject's pre-test score on this measure would continue an excessive degree of error. To overcome this possibility subjects sitting fewer than three internal exams were considered to have missing data on the pre-treatment score of this variable.

4.4 General Anxiety

Four measures of general anxiety were employed in this study. Three of them are derived from the Cognitive-Somatic Anxiety Questionnaire (CSAQ) (Schwartz, Davidson, & Goleman, 1978): the cognitive scale (CSAQ.C), the somatic scale (CSAQ.S), and their sum (CSAQ). The fourth is derived from the State-Trait Anxiety Inventory (STAI) (Spielberger, Gorsuch, & Lushene, 1970): the trait scale (A-Trait).

Noteworthy is that all four are measures of trait general anxiety, that is they measure enduring characteristics of personality and are not situationally specific. Moreover, as pointed out in section 3.6 both the CSAQ.C and the CSAQ.S have yielded significant correlations ($p < .001$) with the A - Trait scale of .67 and .40 respectively (Schwartz et al., 1978).

4.4.1 The Cognitive-Somatic Anxiety Questionnaire

The CSAQ was developed by Schwartz et al. to measure

anxiety in its cognitive and somatic manifestations. The 14 items (seven for each scale) that make up the inventory were selected from "well known questionnaires". They were then "unanimously agreed upon" by three judges as to their reflection of cognitive or somatic anxiety. The items forming the two scales appear randomly intermingled; subjects are asked to "rate the degree to which you generally or typically experience this symptom when you are feeling anxious" on a 5-point scale ranging from "not at all" (1) to "very much so" (5). The possible range of scores for either the CSAQ.C or the CSAQ.S is 7 - 35, and 14 - 70 for the total CSAQ.

The authors of the CSAQ provided some data on its discriminant validity: the CSAQ scales significantly discriminated a group of exercisers from a group of meditators, yet the two groups total CSAQ score did not differ significantly. Although no reliability data was offered by Schwartz et al. (1978), Delmonte and Ryan (1983) reported a Cronbach's Alpha of .85 and .81 for the CSAQ.C and the CSAQ.s respectively. Reported between the two scales have been .42 (Schwartz et al., 1979) and .64 (Delmonte & Ryan, 1983).

The CSAQ employed in this study can be found in Appendix E.

4.4.2 The State Trait Anxiety Inventory

Although a measure of general anxiety, the A-trait has been used to measure test anxiety (c.f. Spielberger et al., 1970); as explained in the previous chapter (section 3.5), this is not how it was intended to be employed here. It was

perceived as a personality variable unaffected by treatment which moderated the effects of treatment on the various outcome measures.

In addition to the A-Trait the STAI contains a scale measuring state anxiety (A-State). While the A-Trait was developed to measure an enduring aspect of personality, the A-State measures a transitory form of anxiety influenced by environmental factors. The degree of correlation between A-Trait and A-State reported in the manual ranges from .11 to .67. Larger correlations between the two scales are obtained under conditions which either pose a threat to self-esteem or involve personal adequacy evaluations. Male samples yield higher such correlation than females.

The STAI has been a very popular research instrument and it does not require a detailed description here. Each of the two scales consists of 20 items with the same "almost never" to "almost always" range. The range of possible scores for either scale is 20 - 80. The manual has norms for college freshmen, undergraduates, high school students, neuropsychiatric patients, general medical and surgical patients, and prisoners. It also reports data on reliability: alpha coefficients range between .86 and .92 for A-Trait and between .83 and .92 for A-State; test-retest reliabilities range between .76 and .84 for the A-Trait and between .16 and .54 for the A-State. Low test reliabilities for A-State are to be expected as it is intended to measure transitory anxiety states.

The STAI has undergone impressive validation procedures (c.f. Spielberger et al., 1970).

4.5 Achievement Motivation

Only one of the variables measured in this study falls under this parameter: achievement motivation measured by the short form of the Ray Achievement Motivation (RAM) scale (Ray, 1979).

The RAM scale employed in this study consists of 14 items with a three response option: "Yes" (scored 3), "Don't know" (scored 2), and "No" (scored 1). Half the items are reverse scored (e.g., "1" becomes "3"). The possible range of scores is 14 - 42.

The RAM has acceptable validity (c.f. Ray, 1979) and good reliability. Alpha coefficients ranging between .72 and .79 have been obtained in samples from Australia and Britain; the South Africa sample alpha was .67.

It ought to be pointed out that this scale is largely a measure of work or career achievement motivation.

The RAM employed in this study can be found in Appendix F.

4.6 Rigidity

Roger Brown (1965) elegantly illustrates this concept.

> Rigidity is a term from common parlance with a root meaning that makes reference to the physical world. To produce changes of form in a substance a degree of resistance must be overcome. When this resistance exceeds our expectations-when a joint moves stiffly or a lump of clay is not malleable-we are likely to call the substance "rigid". Abstracting from the physical case we attribute rigidity to applied

> forces. An elderly person who cannot change his ideas with the changing times manifests rigidity; a patient in pyschotherapy who does not relinquish his defences, in spite of the therapist's insightful interpretations of them, manifests rigidity. The prejudiced person is supposed to show rigidity in his refusal to give up ethnic stereotypes which are presumably contradicted by common experience (p. 506).

In the "Authoritarian Personality (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1969) it was argued that prejudiced subjects tended to manifest rigidity, furthermore, Adorno et al. tended to equate intolerance of ambiguity with rigidity. Frankel-Brunswick (1949) found that prejudiced children were intolerant of ambiguity.

Rokeach (1948) found that children scoring extremely high on ethnocentrism were significantly more rigid in their solutions to the Einstellung Water Jug (EWJ) Test. He went on to argue that the EWJ test is an index of generalized mental rigidity, which stimulated considerable research (e.g., Applezwig, 1954; R. Brown, 1953; Coulter, 1953; in R. Brown, 1965; Goodstein, 1953) and controversy (c.f. Luchins, 1949; Rokeach, 1949).

Roger Brown (1965) pointed out that a comprehensive program of study would be necessary to establish the existence and character of rigidity as a cognitive style and complained that the studies available were too few and heterogeneous. Several diverse measures have been shown in a particular case to correlate with prejudice or ethnocentrism. It is not clear that these various performances operationalize the same concept. Applezwig (1954) reported data suggesting that subjects who are rigid on one measure of rigidity are not

necessarily rigid on others. Correlations between six measures of rigidity yielded only three significant correlations out of 15. Moreover, the three correlations were low and two of them were negative.

Roger Brown (1953) found a relationship between the EWJ test and authoritarianism only under stressful conditions and called for a situation-specific approach to rigidity. Nonetheless, Levitt's (1956) review did not support Brown's hypothesis.

As I pointed out in the previous chapter (section 3.6.4) I chose two diverse measures of rigidity which failed to correlate (Applezwig, 1954; Goodstein, 1953): the EWJ test (Rokeach, 1948) and the Ethnocentrism scale (Adorno et al., 1969). I reasoned that whatever they did measure it was more likely that either of the two would measure rigidity.

For simplicity of exposition I have been referring to the concept of rigidity as if it was an established construct. However, as pointed out above, I do acknowledge that its validity as a distinctive cognitive style remains to be ascertained.

4.6.1 The Einstellung Water Jug Test

There are a number of EWJ tests appearing in the literature (c.f. Levitt, 1956), each of them coming with a slightly different name. The version employed in this study (see Appendix G) was the same as Rokeach's (1948) which in turn was adapted from Luchins (1942). It involves ten arithmetical problems in which subjects are required to

determine how they could measure out various quantities of water using bottles of specified sizes. The best solution is the shortest possible method. An example is given in the EWJ test:

> Given : Containers of capacities: 31, 61, and 4 pints
> Obtain: 22 pints.
> Solution: fill the bottle that holds 61 pints, from it fill the 31 pint bottle, from the remainder withdraw 4 pints twice.
> In short: 61-31-4-4 = 22

The first five problems can all be solved by the same method, which may be abstractly described as: Largest - Second Largest - Smallest (twice). These are the "set" problems. Subjects normally see that there is a "formula" which handles all problems and are pleased to have found it (R. Brown, 1965). However, for the last five problems, while the "formula" still applies a shorter solution is available. For example:

> Given: 23, 49, and 3 pints
> Obtain: 20 pints
> Solution: 49-23-3-3 = 20 or 23-3 = 20

(Note that the containers' capacities are always given in the same order with respect to their magnitude: Second Largest, Largest, Smallest). Many subjects will continue to use the "formula" to solve the last five problems: this is the Einstellung effect or "set". Rigidity on task means perseverence with one mode of thought in the face of information that clearly calls for a new mode. Most researchers using the EWJ test imposed a time restriction of 2 1/2 minutes to complete it; however Levitt and Zelen (1953) applied no such time constraints and found a positive relationship between the EWJ test and ethnocentrism.

In the version employed here subjects were asked (1) to complete it "as quickly as possible" and (2) to record the exact time at the beginning and at the finish. The idea was to get them to work quickly but without time constraints. Rokeach (1948) included a control problem (the second on the list) to which both the "formula" and the short solution would apply - those subjects who would apply the "formula" in solving it would be eliminated from the study. Although the same problems were used in the EWJ test employed in this study the control problem was not used, so as to avoid the risk of eliminating too many subjects from an already limited group size.

Each of the last five problems is scored 1 if "formula" is used and zero if the best (short) solution is applied instead. The possible range of scores is, therefore, 0-5.

Levitt (1956) argued that, because of its nature, no reliability can be calculated on the EWJ test and that its validity as a measure of rigidity (8 years after Rokeach proposed it) was still lacking.

Levitt (1956) reported the distribution of EWJ test scores which is non-normal and U-shaped and concluded that the EWJ test is a poor psychological test qua test. Notwithstanding these shortcomings there is no reason why the EWJ test cannot be found to be predicative of test anxiety treatment outcome and that's how its use was intended in this study.

4.6.2 The Ethnocentrism Scale

The ethnocentrism scale was one of a number of measures developed by Adorno et al (1969) for their studies on authoritarianism. Because the Ethnocentrism scale was developed on the white American population it had to be modified to suit this study's subjects (white New Zealanders). For instance this involve replacing "Negroes" with "Maoris", "CIA" with "SIS". In addition, because the original Ethnocentrism scale was employed as a research instrument some 40 years ago, several items were either modified to suit the current times or were eliminated altogether.

Twelve items (2, 8, 23, 27, 20, 21, 23, 25, 26, 27, 28, 29) were eliminated as they could not be so adapted or because they were obsolete. Twenty-two items, modified or not, were kept to form this new scale which I called Ethnocentrism-New Zealand or E-NZ for short, a copy of which can be found in Appendix H. The response format was kept identical to the original measure which is a 6-point Likert scale. Subjects are asked to mark each statement according to their degree of agreement or disagreement with it by following the scale below:

+1: slight support, agreement
+2: moderate support, agreement
+3: strong support, agreement

-1: slight opposition, disagreement
-2: moderate opposition, disagreement
-3: strong opposition, disagreement

Notice that no zero position of neutrality is provided. Therefore the 22-item E-NZ score can range from -66 to +66. No psychometric properties are available for the E-NZ, but

they are likely to be similar to the original Ethnocentrism scale. Reported in the "Authoritarian Personality" are split-half reliabilities for the Ethnocentrism scale of .91, this scale was also reported to correlate .80 with the original 52-item Anti-Semitism scale suggesting that antagonism to the culturally unlike is a generalized sentiment.

The absence of reverse scored items in the Ethnocentrism scale as well as the Potentiality for Fascism (F) scale (Adorno et al., 1969) and similar scales discussed in the "Authoritarian Personality" has attracted a great deal of criticism (e.g., Bass, 1955; but see R. Brown, 1965, for a balanced discussion on this issue). Acquiescence, not authoritarianism (or ethnocentrism), the critics argued, is what is measured by a variety of scales described in the "Authoritarian Personality". Nonetheless, Heaven (1983) found that acquiescence generalizes across items in a given scale but not necessarily across scales. Further, they found that acquiescers to F scale type items, regardless of content, do exist as a group and that they tend to be similar tolike the people which the F scale tended to identify: conformist, prejudiced, economic radicals, and poorly educated - Authoritarianism research was saved!

4.7 Social Validity

Two questionnaires measured variables falling under the social validity (Wolf, 1978) parameter: the First Impression of Treatment (FIT) questionnaire and the Participant Reactions

Questionnaire (PRQ)

4.7.1 The First Impression of Treatment questionnaire

The FIT questionnaire was administered at the end of the first treatment session for all four groups and it aimed at measuring subjects' expectations for improvement after the introduction of their respective treatment programs. The FIT contained three questions. The first tapped subjects' perception of the program and its expected benefits on the basis of information obtained during that session. The second question asked whether they expected improvement and the third whether they would recommend participation in "one such program" to a test-anxious friend.

No scale was provided and subjects' answers to the second question were scored on the following 5-point scale,

No	Don't	Don't know	Maybe/probably	Yes
-20	-10	0	+10	+20

For the third question the same scale was employed except that -2, -2, +1 and +2 replaced -20, -10, +10, and +20 respectively.

The first question was assessed to decide whether perceptions and expectations about the program at that stage were positive or negative. A negative answer to the first question would automatically cause a -20 score for the second question and a -2 score to the third. All subjects gave a

positive answer to the first probe.

A copy of the FIT can be found in Appendix I

4.7.2 The Participant Reactions Questionnaire

The PRQ included a number of important questions pertaining to the social validity of the four treatments. It contained questions asking whether and how much the "program" and the therapist helped them "in managing your exam anxiety". Subjects were also required to rate the therapist expertise and empathy towards them ("rate the degree to which the therapist has understood your particular problem").

It must be pointed out that in this case empathy would also reflect subjects' acceptance of the treatment and its rationale; after all they were told at the beginning of the program and thereafter how test anxiety comes about, and treatment reflected that conception of it. The validity of treatment and its conception of test anxiety were never questioned apart from one of this groups's subjects at the beginning of the first session of the SES program. However he was quickly persuaded to accept it. Moreover, during a session of the Placebo treatment one subject questioned why the cognitive control scenes (c.f. section 5.7) were not related to examinations (as one would reasonably expect) but even the validity of this treatment was not questioned.

Subjects also had to rate how "authentic" the treatment was to them and express whether they would recommend it to a friend suffering from test anxiety. Appended on the PRQ was a

question probing their opinion of "this program" and to suggest ways on how to improve it.

As the first five questions of the PRQ are answered on a 7-point scale while the last was identical to and scored identically the third questions of the FIT questionnaire (i.e., Yes = 2, Maybe = 1, Don't know = 0, Don't = -1, No = -2) the possible range of scores on the PRQ is -2 to 37.

A copy of the PRQ can be found in Appendix J.

CHAPTER V

METHOD

5.1 Subjects

Thirty seven subjects (30 females, 7 males) aged between approximately 18 and 50 years took part in this study. They were recruited on the basis of:

(a) voluntarily responding to an announcement made at the university lecture rooms for subjects for a study on test anxiety "the aims of which are to help students manage their anxiety during exams and, hopefully, improve their grades"; and

(b) yielding a score of above 32 in the Debilitating Scale of the Achievement Anxiety Test. This cut-off score has been shown to include approximately the upper 13% of the distribution of scores on this measure (c.f. section 3.1.4).

The subjects were recruited during the first week of the third term, mostly from first year courses in psychology and education. I would make arrangements with the lecturer concerned to approach the students either at the beginning or at the end of their lecture.

The following announcement was made to the students in the lecture room: "Test anxiety!... As you probably know a lot of students feel very anxious during exams. As well as being discomforting , such anxiety negatively affects your grades. The aim of this study is to enable you to manage your anxiety

during tests and exams and, hopefully, to improve your grades... I have a questionnaire for those interested in participating, it only takes 5 minutes". The AAT was then distributed to the interested students who completed it at their seats and handed it back on their way out.

They were asked to write their name and telephone number on the questionnaires so that they could be further contacted. Students who scored above 32 in the Debilitating Scale (AAT-) were contacted by phone and offered treatment. The following script was used with each student over the telephone: "On the basis of your questionnaire results you have been selected to participate in this study". Those who agreed to participate were then told, "You'll be meeting twice a week with a group of about ten people for 50 minutes". At this point some students declined the offer, as they could not afford the time.

Those who still agreed to participate were informed about the meeting time and place of the group they had been assigned to. Each student was assigned to the group with the fewest subjects at the time of contact. In the instance of clashes with the subject's own schedule a choice of another group was offered.

It was stressed to all of them that it was very important that they complete the program if they chose to participate. Forty-four subjects were so recruited.

Holroyd's Cognitive-Attentional Training (CATH) and CATH plus Study and Exam Skills (PK) groups had 12 subjects each, the Study and Exams Skills (SES) group 11, and the Placebo (P) group nine. Following the first treatment session one subject dropped out of the PK group which was left with 8. A subject

from the SES group also discontinued her attendance after the first session and one other after the second, thus leaving this group with 9. After the sixth session one subject from the CATH group withdrew, but the remaining 11 completed the program. The only male subject in the Package (PK) group withdrew after the second session and one other could not complete the program because of illness. Failing to complete the questionnaires appropriately or by the specified time, a third subject from this group had to be excluded from the study, thus leaving this group with nine for the statistical analysis. The four treatment groups had 37 subjects in total completing both their respective program and the questionnaires.

A further 22 subjects who scored above the AAT- criterion were offered treatment, but for various reasons declined to participate. These served as a control group for one of the academic measures analyses. These subjects were no longer contacted as final grades are available from university records.

5.2 Procedure

Having already completed the AAT those subjects initially recruited were given the Inventory of Test Anxiety (ITA), the Cognitive-Somatic Anxiety Questionnaire (CSAQ), and the Inventory of Study Habits (ISH) to complete before meeting with their respective groups.

The groups met twice weekly for 4 weeks. During the first meeting subjects did not receive treatment but completed form

1B of the School and College Achievement Test II (SCATT), which takes about 45 minutes, and the Test Anxiety Self-Efficacy Questionnaire (TASEQ). At the end of the following meeting when treatment began, subjects in each group were given the First Impression of Treatment (FIT) questionnaire to complete. At the end of the last treatment session subjects were given the Participant Reaction Questionnaire (PRQ) to complete. They were also given the ITA, the AAT, the ISH, and the TASEQ to complete "as soon as you can". In any case they were instructed to complete them "before exams begin". Most subjects filled out the questionnaires during the next couple of days, some took a little longer, but all completed it within one week and well before exams began. Furthermore, they were given a form to complete and sign, giving the author consent to look at their final exam marks as well as their term tests and exam marks.

Subjects were subsequently asked to collect two copies each of the TASEQ, the Post Test Anxiety Self-Efficacy Questionnaire (P-TASEQ) and ITA. They were to be completed as follows: the TASEQ before their first exam, and the P-TASEQ and the ITA after that exam. The same procedure was to be followed for their last exam. Only 21 subjects completed these questionnaires satisfactorily, i.e. completed all six questionnaires appropriately and in the desired sequence.

Four weeks later, once exams were over (and subjects had more time at hand) they were given nine questionnaires to complete, the AAT, the ITA, the TASEQ, the ISH, the PRQ, the State Trait Anxiety Inventory (STAI), the Ethnocentrism-New Zealand (E-NZ), the Einstellung Water Jug (EWJ) test, and a 14-item short form of the Ray Achievement Motivation (RAM)

scale. All but one subject completed the 4-week follow-up questionnaires; the subject who did not belonged to the PK group.

Twelve months later (the follow-up 12) subjects were asked to complete the following questionnaires; the AAT, the ITA, the TASEQ, and the two scales of the ISH relating to exams - Studying for Examinations (SE) and Examination Technique (ET). They were also asked to complete the SCATT (Form 1a) which was done in the same room where treatment sessions were held.

Final grades were obtained in person from university records but tests and exam marks had to be obtained from each department concerned.

5.3 Treatment

Each group met twice weekly in the activity room of the Psychology Department for about 50 minutes each time. Occasionally a session would last a little longer more often with the PK group but never with the SES group. The SES group, and the CATH group met at 4 pm (on Mondays and Wednesdays, and Tuesdays and Thursdays, respectively) while the P group and the PK group met at 5 pm (on Mondays and Wednesdays, and Tuesdays and Thursdays respectively). All discussions and treatments happened around a table except for the P group where, lacking chairs comfortable enough to relax on, subjects shifted their chairs near to the wall so they could rest their heads against it. As this treatment progressed, following a request by one subject to relax lying down on the carpeted floor, all subjects carried out the

meditation exercise in this fashion.

I served as therapist for all groups and tried to follow each manual with the same care and accuracy. I made every effort to show the same interest, enthusiasm and empathy towards all subjects irrespective of their group. I felt uneasy with the P group (because of the belief that it was an inert procedure), particularly during the first session or two, but I don't think I let it show.

In the event of a subject missing a meeting a catch-up session would be done with that subject alone, except for the P group where a catch-up session was held for three subjects at once.

I shall now describe in some detail each of the four treatments and how they were applied in their respective sessions.

5.4 Holroyd's Cognitive-Attentional Training

This treatment closely follows the manual used by Holroyd (1976) in his study. It is based on Wine's (1971) cognitive-attentional model of test anxiety. According to this view the debilitated performance and subjective discomfort of test anxious individuals is due to the maladaptive cognitive responses that are elicited by evaluative situations which distract attention from the task at hand and interfere with performance.

Heightened autonomic responses to the test situation are assumed to exert a minor influence on the test-anxious persons behaviour: it is not the intense autonomic responses per se

which trouble the test anxious individuals but the labeling of them and the directing of attention to them. Indeed it is hypothesized that it is such labelling and direction of a attention toward the specific manifestations of arousal and away from the task at hand which disrupts the student's test behaviour. It is these two factors which distinguish a highly test anxious individual from his/her counterparts.

This treatment focused exclusively on eliminating the test anxious subject's task-irrelevant thoughts and on directing attention to the task at hand, whether it was studying for exams, or answering exam questions.

The specific therapeutic procedures were originally derived from Ellis's (1962) rational emotive therapy.

The CATH rationale emphasized that group members could learn to manage their test anxiety by learning to control task irrelevant thoughts that generate anxiety and distract attention from their work. Initially subjects were encouraged to identify the context in which they appear, and the negative behavioural consequences of these thoughts. The therapist questioned the logic and the validity of these (negative) statements and encouraged clients to replace such thoughts with alternate (positive) self-instructions designed to avoid anxiety arousal and to direct attention to the task at hand. Persuaded by the merits of treatment rationale, subjects were given a role-play exercise, some covert practice and some direct practice, to help them master the habit of using positive self-instructions and to direct their attention to the task at hand while sitting a test or an exam.

5.4.1 The seven CATH treatment sessions

Here is a session-by-session description of this treatment. TABLE 5.4.1.1 contains a synopsis of this treatment procedure. Each session (except the first and last) began with a 5 to 10-minute discussion. This served to break the ice and to enhance rapport; during this time subjects were asked to raise any questions or problems they might have had and to point out any new negative self-statements they became aware of. The therapist corrected any misconceptions concerning treatment and briefly repeated the rationale for this treatment. A brief 5 minute discussion concluded each session. Here the therapist elicited any reactions to the day's session, answered any queries, and reminded subjects to carry out their "evening review".

Insert Table 5.4.1.1 here

This review involved the silent recollection of the day's events; its aim was to identify any negative self-statements the subject had made at any time throughout the day. Subjects were instructed to carry out this review with their eyes closed and to begin it with a few deep breaths; they were also to construct a list of negative self-statements and to add to it any new ones they became aware of. Particularly during the last three sessions the following question was raised to the subjects; "How is what we are learning here helpful in reducing test anxiety?"

TABLE 5.4.1.1

Synopsis of Holroyd's Cognitive-Attentional Training procedure

Session 1:	
	(a) Personal introduction and statement of nature duration and extent of test anxiety for each client (10 min)
	(b) Explanation of rationale and course of treatment (5 min)
	(c) Instructed analysis of thoughts during test situations (30 min)
	(d) Group discussion and homework assignment (5 min)
Session 2:	
	(a) Discussion of homework assignment and correction of any misconseptions of treatment (10 min)
	(b) Construction of a list of anxiety arousing self-statements by group discussion (15 min)
	(c) Instruction in the use of alternate self-statements (20 min)
	(d) Group discussion of session (5 min)
Session 3 and 4:	
	(a) Group discussion of problem areas including new anxiety arousing self-statements (10 min)
	(b) Individual client-therapist interaction: training in the use of alternate self-statements (30-35 min)
	(c) Group discussion of session (5 min)
Session 5:	
	(a) Group discussion of problem areas including new anxiety arousing self-statements (10 min)
	(b) Role play exercise (30-35 min)
	(c) Group discussion of session (5 min)
Session 6:	
	(a) Group discussion of problem areas including new anxiety arousing self-statements (10 min)
	(b) Role play exercise as in Session 5 but using different negative self-statements from client's own list (15 min)
	(c) Covert practice in the use of coping skills learned in previous sessions (15-20 min)
	(d) Group discussion of session
Session 7:	
	(a) Test practice of skills learned (20 min)
	(b) Group discussion of test practice (20 min)
	(c) Group discussion of program (10 min)

Session 1

- This session began with each subject's personal introduction and a brief description of the nature, duration, and extent of test anxiety. Moreover, they were to tell whether they experienced anxiety in other social or evaluative situations.

- The therapist then went on to explain the rationale and the course of treatment.

- Next, subjects were asked to carry out an analysis of their cognitive and attentional processes during test situations. They were requested to close their eyes and imagine an exam scene; they were then instructed to recall the kinds of thoughts and feelings they had experienced in that situation, "Imagine you are watching a movie but watch the events as they flow by on a mental screen. What do you find yourself saying to yourself and thinking? What are you worried about, at what point does this occur to you? When do you start feeling anxious?" This review was continued for about 30 minutes.

- Following this the therapist encouraged a brief discussion of the group's experiences, emphasizing the part played by subjects' negative self-statements in test anxiety. The "evening review" assignment was described and subjects were asked to spend no more than 15 minutes each day doing it. They were also requested to construct their own list of negative self-statements and to bring it along next time, as they were "going to work from it".

Session 2

- This began with a discussion of the homework assignment and correction of any misconception of treatment.

- About 15 minutes were then devoted to the construction of a list of anxiety arousing and distracting thoughts characteristic of each subject. Group members were encouraged to discuss those distracting thoughts which typified their anxiety and to specify the context in which they occurred. Each subject was encouraged to make a list of the self-statements that occurred most frequently and were personally most debilitating. The list was to include self-statements which occurred while studying, preparing for and actually taking an exam.

- Following this discussion the therapist provided group members with an illustrative list of anxiety coping self-statements. These statements illustrated the type of coping self-instructions that can be useful at each of the three phases of exam taking: (1) confronting and handling the initial stress of the exam (e.g., what is it I have to do? No negative self- statements. Just think rationally. Focus on the task: exactly what does the question really ask?) (2) Coping with runaway anxiety or feelings of helplessness when they do occur (e.g., don't get anxious: just take a moment off and focus my attention on what I have to do. This is the anxiety that I thought I might feel. It's a reminder to cope. Slow down a little. Don't rush and get all in a panic ... there is time for most of it. Label my anxiety from 0 to 10 and watch it change. Now I am under control - back to the exam). (3) Recognizing and reinforcing the successful use of these coping strategies (e.g., It's working. I can control it. I feel I am in control. I made more out of my fear than it was worth. My damn ideas; that's the problem. When I control them I control my fear. I did it!). Subjects were

encouraged to select coping self-statements from the list provided, or preferably devise their own anxiety-reducing, task-oriented self-instructions. It was emphasized that becoming aware of anxiety-arousing and distracting thoughts (negative self-statements) and replacing them with task-oriented and anxiety-reducing self-instructions (positive self-statements), such as those provided, is an effective method for reducing anxiety.

- Each group member was encouraged to construct his or her own list of anxiety reducing and task oriented instructions.

- The session ended with a discussion of the material covered in that session.

Session 3

-To begin, there was group discussion of problem areas including new anxiety-arousing self-statements.

-Then followed individual client therapist intervention: anxiety-coping training. during these 30-35 minutes the therapist tried to identify the habitual thinking style that contributed to each group member's anxiety, and tried to change their cognitive style. Initially he encouraged the subject to identify the content of disruptive thoughts s/he succumbed to, the environmental context within which these thoughts occurred, and to identify their negative behavioural consequences. The therapist also tried to identify the irrational beliefs implicit or explicit in each subject's self talk. The therapist then challenged those anxiety engendering thoughts and the subject's irrational belief while encouraging him/her to use techniques for coping with them. Specifically the therapist questioned the logic and validity of those statements and beliefs and encouraged the practice of

replacing those disruptive thoughts with alternate self-instructions designed to facilitate task attending. As a homework assignment subjects were asked to add to their earlier list of negative self-statements a list of positive and anxiety-coping ones. Specifically they were asked to change each of those negative self-statements into a positive self-instruction and to write it down on the right-hand side of the page (the negative self-statement was on the left of the page). The list was to be handed to the therapist (during the next session) who would read it and put his comments on it.

- This session ended with a discussion of the material covered during that hour.

Session 4

-This was the same as session three except that the list with the negative and corresponding positive self-statements was handed to the therapist.

Session 5

- Group discussion was held, dealing with problem areas including new anxiety arousing self-statements.

-The therapist handed back lists to subjects who read his comments and his suggested changes.

-The therapist encouraged reactions to his comments and to any changes he had suggested for each subject's list. It was made clear that these were genuine suggestions with no mandatory overtones.

- Role play exercise. During these 30 minutes each group member swapped role with the therapist who played the role of a test anxious student bothered by one or two negative self-statements taken from the subject's own list. The

exercise consisted of two phases. Firstly the 'therapist' had to show the 'test anxious student' that his or her negative self-statements were irrational, and secondly the 'therapist' had to help the 'test anxious student' to change such statements into positive ones. I would hold cards with instructions on them to help the 'therapist' in his/her role.

- There followed the group discussion of this session.

Session 6

-This began with group discussion of problem areas including new anxiety- arousing self-statements.

- Role play exercise as for session 5 for 15 minutes. Negative self-statements different from those used in the previous session but still from the subject's own list were used.

- Covert practice (15 minutes). The therapist would ask subjects to close their eyes and recall an exam scene in a similar fashion to that of session 1. Subjects were asked to think of the negative self-statements that would normally come to mind but this time to replace them with positive ones as they had learned during previous sessions.

- The usual end-of-session group discussion was held.

- The group was reminded that during the next session a short test would be held.

Session 7

- Test practice. Subjects were given 12 minutes to complete the Wonderlic Personnel Test. It was emphasized that they should put into practice the skills they had learned during the program to control their anxiety. At the end of the test subjects were asked to write down how anxious they felt, whether they had any disruptive thoughts and whether they felt

they had controlled their anxiety and negative self-talk during the test.

- Discussion on subjects' reaction to the test and their ability to control test anxiety was held. More suggestions were given to the few who still experienced some difficulty (most subjects coped well during the test).

- The usual end of session discussion was followed by an overall discussion on the treatment and the subjects' reactions to it.

5.5 Study and Examination Skills

This treatment aimed at improving study skills with particular emphasis on preparing for and tackling examination questions. It was based mainly on "Effective Study" by Robinson (1970) and the 'Study Habits Inventory and Evaluation Kit' (SHEIK) by Jackson, Reid, and Croft (1979) with additions from texts by Rygor and Wark (1971) and Deese and Deese (1979); moreover, it contained guidelines suggested by Fox (1967) and Goldiamond (1967).

The rationale offered for this treatment is that test anxiety is the result of lack of confidence or knowledge of exam material together with poor examination techniques and that it is only by perfecting one's own study skills and examination-taking skills that test anxiety can be brought under control. A student equipped with good study and examination-taking skills would effectively learn exam material and would know how to tackle exam questions with confidence and without anxiety.

The SHEIK contains seven units: Place of Study (1), Study Times (2), Organization for Study (3), Textbook reading skills (4), Taking notes (5), Preparing for Examination (6), and Taking Examinations (7). Each has instructions on improving study skills and an exercise testing the material learned in that unit. Each of the seven units was photocopied and made into a small handout, a copy of which was given to each of the subjects in this group during the various sessions of the program. Students were asked to read the handout and to complete the exercise at the end of it as a homework assignment. The completed exercise would be handed to the therapist who would bring it back at the following sessions marked and with comments on it where required.

5.5.1 The seven SES treatment sessions.

Group discussions were held at the beginning (except the first and last) and at the end of each session when the therapist answered any queries subjects might have had and encouraged comments and questions on the material being covered. The initial discussion was used as an opportunity to break the ice and to enhance rapport. It was emphasized at different times during the program that the guidelines offered were based on sound research and that they were to be treated as guidelines: If subjects felt uncomfortable about following them they ought to change them to suit their style. Subjects were encouraged to take notes on the material covered during each session.

Here is a session by session account of this treatment.

Table 5.5.1.1 includes a synopsis of this program. Care was taken not to include aspects of the CATH in this treatment, something which was so easy as to be tempting.

Insert Table 5.5.1.1 here

Session 1

- Personal introduction was made, followed by a statement of the nature, duration and extent of test anxiety for eachsubject.

- A statement of the rationale and content of the program was made.

- Place of study: its importance for effective study was emphasized. Guidelines were offered as to the desk and on the lighting, ventilation, and heating of the place of study. Suggestions were also given on how to minimize distractions. Subjects were encouraged to use, as far as possible, the same place for working and to use it for nothing but study.

- Daydreaming technique. To deal with problematic daydreaming typical of several subjects, they were instructed to (1) decide to leave the desk each time they found themselves doing so, but (2) read one page of the material they were working on, and (3) make sure that they left the desk then even if they no longer daydreamed after reading that page and even if they felt like carrying on, working or reading. Details of this technique are offered by Fox (1967). They were asked to use this technique for the following 5 days.

TABLE 5.5.1.1
Synopsis of Study and Examination Skills procedure

Session 1:
- (a) Personal introduction and statement of nature, duration and extent of test anxiety for each client (10 min)
- (b) Explanation of rationale and course of program (5 min)
- (c) Place of study (15 min)
- (d) Daydreaming technique (10 min)
- (e) SQ3R: Survey only (5 min)
- (f) Group discussion of session and homework assignment (Unit 1 of SHEIK) (5 min)

Session 2:
- (a) Group discussion of problem areas (10 min)
- (b) Study times and plan for study (20 min)
- (c) More on daydreaming technique plus Goldiamond (1967) suggestions (5-10 min)
- (d) SQ3R: Questioning and Reading (10 min)
- (e) Group discussion of session and homework assignment (Unit 2 of SHEIK) (5 min)

Session 3:
- (a) Group discussion of problem areas and comments on the study plan to each client (15 min)
- (b) Note taking (20 min)
- (c) SQ3R: Recitation and review (10 min)
- (d) Group discussion and homework assignment (Unit 5 of SHEIK) (5 min)

Session 4:
- (a) Group discussion of problem areas (10 min)
- (b) Reading skills: skimming, speed reading, critical reading, active browsing (20 min)
- (c) More on SQ3R; SQ3R underlying (15 min)
- (d) Group discussion of session and homework assignment (Unit 4 of SHEIK) (5 min)

Session 5:
- (a) Group discussion of problem areas (10 min)
- (b) Exam preparations (30-35 min)
- (c) Group discussion of session and homework assignment (Unit 5 of SHEIK) (5 min)

TABLE 5.5.1.1 contd.

Session 6:

(a) Group discussion of problem areas (5 min)
(b) A little more on exam preparation (5 min)
(c) Techniques for answering essay type questions (20 min)
(d) General exam techniques (15 min)
(e) Group discussion of session and homework assignment (Unit 7 of SHEIK) (5 min)

Session 7:

(a) Test on material learned in the program (15-20 min)
(b) Discussion of test, anxiety experienced and other problems (10 min)
(c) Objective examination questions (10-15 min)
(d) Discussion of session and about program in general (10 min)

- SQ3R: Survey only. Subjects were introduced to the SQ3R and the first step (surveying) was described. Furthermore, they were encouraged to apply this step of the SQ3R each time they studied from textbooks till the next meeting (5 days later).

- A group discussion of the session was held and the homework assignment was given. Unit 1 of the SHEIK was handed out.

Session 2

- Group discussion was held on problem areas and for correction of any misconception of treatment.

- The common problem of students deciding to do some study yet not actually sitting down at the desk to do it was outlined, together with the stresses associated with it.

- Study times and plan for study. Subjects were encouraged to follow several guidelines relating to study times and the plan for study (e.g., study when fresh; take a 5-minute break after about half an hour of study; take a half-hour break after 2 hours of study; avoid working late at night; and make a habit of studying in the morning; make use of the vacant hours between classes). Subjects were provided with a sheet (with days of the week and hours of the day) and were encouraged to devise a weekly schedule of work and leisure. Suggestions were given on how to carry out this task (e.g., enter things already set first, eating, sleeping, classes and lectures, outside work, sport; be specific on the type of work to be scheduled at a particular time; be realistic in doing such assignments, i.e., plan to complete a certain amount of work that can be done in that period. At the same time it was stressed not to be too lenient on such scheduling since too little would then be accomplished). Subjects were to schedule

for leisure times in their study plan and were encouraged to relax and enjoy themselves during such designated periods without feeling guilty (an oft experienced feeling of students who even during times of leisure feel they ought to be studying - which of course takes much of the enjoyment and relaxation out of leisure). But it was made clear that during those periods they assigned for work they ought to do just that. Other suggestions given were, (1) Try to finish your work within the time set, (2) Do not waste time deciding what to study: study the material you have scheduled, (3) Set a time or page limit on the work to be done and if you finish early take a break. It was emphasized that this exercise is about developing good habits for studying, not making decisions to get to work. It was stressed that unlike many jobs and activities the job of a university student is unstructured and s/he needs to structure it in order to get the work done; the idea was advanced that they should be their own manager.

- Subjects were asked to construct their week's schedule and to bring it along to the next session when the therapist would check it and give more comments and suggestions.

- Daydreaming technique: not one page but two were to be read before leaving the desk.

- Following Goldiamond (1967) subjects were encouraged to find out the meaning of unfamiliar words each time they encountered them during their study and to find out and learn clearly any concept they came across which they were not clear about.

- SQ3R: Questioning and reading. Subjects were encouraged to carry on with surveying textbook material which was about

to be read and from then on to make up questions at the beginning of each section in a chapter ("changing headings into questions is a conscious effort to orient yourself towards the material discussed in the text") and read it attempting to answer those questions. Active (as opposed to passive) reading was stressed ("reading becomes an active search for the answers to those questions you posed at the beginning", "you should have in mind what you want to learn as you read each section".

- A group discussion of the session followed, homework was assigned and unit 2 and 3 of the SHEIK were handed out.

Session 3

-First there was group discussion of problem areas, checking and commenting on each subject's study plan.

- The common problems of students taking too many notes and copying notes from textbooks were outlined.

- Note taking. Suggestions were given on how to take effective notes: "Keep notes for different subjects separate; keep details to a minimum; include all main points; organizing notes (use an outline format, labeling, indentation, numbering, emphasis marks, use a cue word at the beginning of each note); editing; and reviewing. It was suggested that students should spend more time listening to lecturers and less taking notes. It was emphasized that notes are best and more easily remembered when in one's own words rather than copied from the book.

- SQ3R: Recitation and Review. The importance of recitation in remembering material just read was stressed (it was also pointed out that there is a body of psychological research supporting this fact). Subjects were encouraged to develop

the ability to recite material just read for memorizing. Recitation "is to be carried out at the end of each section in a chapter",..."it is best to jot down notes from memory in one's own words" (avoiding looking in the book), then one should "check for errors of omission and commission by checking with the text just read". Subjects were encouraged to avoid the common mistake of taking notes while reading the text. It was also suggested that students should review each chapter just read or each lesson just covered by recalling the main points first, then opening the book and covering each section under the main point and recalling each subpoint. At the end of each stage of review subjects were encouraged to check for errors of omission and commission.

- It was emphasized that the SQ3R is an effective technique for textbook reading and that, as with most newly acquired skills, one feels uneasy in the early stages, being tempted to slip back to the old habits. However, subjects should practise the skill until it becomes natural to use it, as the long term savings and benefits of the SQ3R were considerable.

- The usual group discussion was held and the homework assigned.

- Unit 5 of the SHEIK was handed out.

Session 4

-Problem areas were discussed by the group, and correction of misconceptions related to material covered to date was made.

- The importance of recitation was re-emphasized and it was pointed out that the initial difficulties of using the SQ3R are far outweighed by its long term results. It was suggested that the amount of study time devoted to recitation should be proportional to the difficulty of the material studied ("as

much as 90% for difficult material rich in detail and confusing relations", "about 50% for subjects like psychology, education, economics, sociology and political science", and "only about 20% for a story or a well organized history passage").

- Various types of reading were described and their usage outlined. Skim reading was introduced as an effective way to (1) preview a book, (2) get the main idea out of a book or chapter or (3) extract important details through active browsing of the text.

- Subjects were also encouraged to be critical and evaluative of the material they read and to expand on it trying to see the relevance of the learned material to social or personal or everyday problems ("This is likely to keep your interest up" and "to link what you learn with what you already know - a golden rule for understanding and remembering").

- Subjects were given instructions on how to improve their reading by building on their vocabulary (by paying attention to new words; using a vocabulary, vocabulary cards; distinguishing between general and technical words) as well as learning to read faster (but "be sure that your comprehension does not suffer", they were cautioned).

- SQ3R underlying. This modification of the basic SQ3R was introduced, the difference being that with this version you do not jot down notes during the recitation stage, but after reciting you underline key words and phrases. The use of a marking system that will show key points and their degree of importance was encouraged. It was emphasized that one ought to carry out the underlining of key words and phrases after recalling the main points (i.e., after recitation of answers

to the questions posed before reading a section).

- Group discussion of session and homework assignment was held.

- Unit 4 of the SHEIK was handed out.

Session 5

- Problem areas were discussed and misconceptions clarified.

- The basic SQ3R was presented as superior to its variations but it was acknowledged that some students may be more inclined to use the SQ3R underlining.

- Preparing for examination. In preparation for each exam subjects were encouraged to: (1) find out details of the exam (what material is being examined; what type of questions were going to be asked, i.e., whether essay type, multiple choice or short answer, or a mixture of them would appear on the exam paper; whether there were going to be choices; how many questions were to be answered; and how long the exam was going to be. (2) use old exam papers (to learn the format that was going to be used, to look for questions that are likely to come up again, to answer questions on such papers and check answers). It is very good practice to review material on questions that have been answered wrongly and material on questions that could not be answered. (3) generate questions and answer them as they could probably generate similar questions to the ones that were going to appear in an exam. (4) concentrate on things that aren't known as it is not necessary to review material already known. (5) make a definite plan for study for the previous week or two before the date of the exam. (6) plan for review periods separate from study periods (7) plan for review periods of no longer than about 1 hour as reviewing is fatiguing. (8) prepare in

the same fashion for essay type and objective (multiple choice, true-false, etc.) exams. (9) spend no longer than about 1 to 2 hours studying the night before the exam (this should be spent reviewing). (10) get enough sleep the night before the exam as it is important to be fresh in order to do well in an exam; not to worry if they could not go to sleep - just relax instead. (11) the suggested methods for review were: doing a test or quiz soon after learning a lesson, and rereading (relearning) the lesson for material learned more than 2 weeks before. Near exams subjects would be likely to study for longer hours, while recreation would be minimized. However they were encouraged to maintain their eating, sleeping, and exercise patterns.

- The usual discussion was held, the homework was assigned and unit 6 of the SHEIK was handed out.

Session 6

- The group discussed problem areas and misconceptions were clarified.

- More on exam preparation. The importance of attending classes before an exam was emphasized as significant information relating to it is often given then. Subjects were also encouraged to:

(1) Summarize the material for an exam into a few pages of notes.(2) Keep on answering exam questions and preferably take an exam similar to the one coming, ahead of time. For those people who were interested in studying in groups a 'round robin' strategy was suggested; this consisted of a member of the group giving an exam question for another group member to answer. The person next to him or her would in turn add to the answer already given, the next person would add to

the answer given to that point, and so on.

- Cramming was introduced not as a substitute for planned study but as useful if for any reason subjects ran out of time. Subjects were told "the best time for cramming is the day before the exam", but on the night before the exam to spend "no more than 1 or 2 hours working, preferably reviewing", then "get some sleep".

- Taking exams. The view that "exams are a cooperative venture between lecturers and students" "trying to find out where effort is needed" "not a battle in which each tries to outwit the other" was stressed, for students should learn to appreciate "all the time and attention" given by their instructors rather than fear them and the exam questions they set.

- Students were encouraged to: (1) Get to the exam room 5 to 10 minutes early, not later, as rushing would be stressful. By the same token they should not arrive much earlier as observing nervous students or talking to friends who are nervous about the exam could also make them nervous. (2) Not to worry about having enough time as the time allotted for the exam has been planned to be sufficient to answer all questions required. (3) Spend 30 seconds or so glancing over the exam paper to get an idea of (a) how long the paper is, (b) whether certain parts carry more marks than others and therefore require more time to be spent on them, and (c) whether certain parts are easier to answer. (4) "Budget your time and stick to it !" - this was stressed as a very important procedure (ways to appropriately divide the time available for each question are described in unit 7 of the SHEIK). (5) "Start on a question you can do more easily first" as this "will put you

on a good footing and get you started more easily". (6) "Do your best"! Exams are made a little difficult to give a range of scores for grading, "so don't worry if you can't answer a question to perfection" - there would be no point in setting an exam where everybody could answer all the questions easily. (7) "Think and write one question at a time", "don't worry about questions further down the list until you get to them". This was stressed as a very important point to remember.

- Examination technique: A common error students make in answering questions, namely interpreting a question as "write everything you know on this topic", was pointed out and subjects were warned against it. It was explained that time constraints lead examiners to ask specific questions. Given that it is customary not to give marks for material which does not answer the question it is very important that one understands what "the question asks exactly before you begin to construct your answer". Subjects were told to imagine a tape recorder in front of them that would keep on playing these words: "Answer the question, answer the question, answer the question ...!" throughout the exam. Subjects were reminded that the first step in answering an exam question is to note the key word and preferably underline it. The second step is to note limiting words. The need to have a clear understanding of words that are often in exam questions (e.g., compare, contrast, discuss, evaluate, etc.) was also stressed. To this end, a two-page list of such words and what they meant was provided for each group member, supplementing the one appearing on unit 7 of the SHEIK. Subjects were advised to answer the question only when they had a clear understanding of it. Subjects were also warned about another common

problem, which is to write the first idea that comes to mind, then write the second idea that comes to mind, and so on, thereby forgetting to keep to the topic. Subjects were encouraged to write a plan of their answer before answering the question. This was also stressed as a very important procedure. To make the plan subjects were instructed to jot down the main points of their answer in outline format taking care that they were in logical order; "this brings other subpoints and ideas to mind that can be inserted in the correct spots in your plan" and "presto! in a few minutes your plan is done"; "once such a plan has been made writing your answer is 'dead easy', merely expanding on those points outlined in the plan." Subjects were reminded to incorporate an introduction and a conclusion when answering essay type questions, and further, to use headings and numberings, where appropriate. Finally, they were to spend 5 to 10 minutes checking their answer. It was pointed out that handwriting is important in answering exam questions; "a lecturer cannot mark an illegible script". In the attempt to improve their own handwriting subjects were encouraged to concentrate on improving their least legible two or three letters, which usually results in remarkable improvements.

- "Taking essay examinations", a section from "How to Study" (Deese and Deese, 1979) was photocopied and handed to each of the subjects.

- The final group discussion of the session was held and subjects were reminded of a small test to be held at the beginning of the next and final session of the program, testing material learned in the previous six sessions. Finally, unit 7 of the SHEIK was handed out.

Session 7

- Fifteen minutes were spent on the test. Subjects were required to answer the following questions, "How to answer essay type questions in exams" (because of its relevance in this context it was made compulsory), and two others from the following,

(1) How a good place of study should be.

(2) Why it is important to write a plan for study and how to do it.

(3) What the best way to take notes is

(4) Write short notes on the SQ3R

(5) Write short notes on study times.

(6) What things should you do when preparing for exams.

The format of the test was similar to that of many exam papers. The aim of this test was twofold; firstly to test their knowledge of important material covered during the program, and secondly to give subjects some practice at sitting a test and getting feedback on their reactions to it.

- A discussion of the test, experienced anxiety, and any other problems subjects experienced followed. The test was to be marked and handed back to them.

- Subjects were encouraged to develop the habit of making use of returned material by going over it, understanding mistakes, and asking instructors if they still have difficulties. This was presented as a good chance to review important material not properly understood, as it was pointed out that "usually only important material is examined".

- Objective examinations: In answering this type of question subjects were encouraged to (1) survey the exam paper, (2) read instructions carefully, (3) be sure to understand the

scoring rules, (4) answer easy questions first, (5) place each question in context when answering it, and finally, (6) check the answers given and change those they feel need to be changed, as research shows that such practice improves grades (c.f. Robinson, 1970).

- True/False questions: Subjects were given a list of qualifiers and it was suggested that they substitute them for the qualifier in the original statement. Where this produces a statement that is better than the original, "false" is the correct answer, otherwise it is "true". They were also encouraged to scrutinize key words when answering such questions as they are likely "to tell you whether it is true or false".

- Multiple choice exam questions. To answer these types of questions subjects were encouraged to look for items which are obviously wrong and cross them out, then concentrate on the "couple of items left" to choose the correct one.

5.6 The Package procedure: CATH plus SES.

This treatment consisted of the combination of the two treatments so far described. The content of session 1 for this treatment consisted of session 1 of CATH treatment plus session 1 of the SES treatment, similarly for session 2, and so on for all seven sessions. Consequently the pace was faster during these sessions.

Attempt was made to cover all material within 1 hour, although sometimes it took a little longer.

The rationale offered for this treatment was that test

anxiety is the result of maladaptive thinking and inappropriate attentional focus together with poor study and examination taking skills which make students feel little confidence in their knowledge of exam material. By learning to control their thought processes and attentional focus and by perfecting their study and examination taking skills, which would make them more confident about taking exams, "test anxiety will be under control".

Like the previously described treatments the pratice test took place at the beginning of session 7. It was 18 minutes long, consisting of an attempt to complete half the Wonderlic Personnel Test, the compulsory question on how to answer essay type exam questions, and the choice of one (as opposed to two) from the remaining six. These questions were the same as those for the SES group. As for the CATH group, shortly after finishing the test subjects were asked to make a note of their feelings and thoughts on the back of the Wonderlic Personnel Test. These were similarly discussed.

5.7 Placebo

This treatment was based entirely on Holroyd's (1976) pseudotherapy attention control procedure. It was included to assess the degree of improvement that might be attributed to nonspecific treatment effects, such as expectation for improvement, attention, warmth, empathy, and interest of the therapist; and implicit demands for improvement generated by impressive therapeutic procedures.

5.7.1 The seven P treatment sessions

The rationale for this treatment was that meditation exercises would allow group members to achieve a mental state that could not be disturbed by test anxiety, i.e., being able to control their mind while still aware of bodily sensations.

The "instructed body awareness training" involved instructing subjects to "settle back as comfortably as possible", close their eyes, take a few minutes to clear their minds "of the cares and thoughts of the world", and try to be "as blank as possible". And if thoughts intruded, "let them run their course and then clear your mind". After about 10 to 15 minutes clients were instructed to concentrate on various sensations of the hands, arms, face, and other parts of the body to increase "feelings awareness". Following the second sessions, subjects were to carry out (a) the "instructed body awareness training", (b)"imagine a commonplace situation such as riding a bicycle on a tree-lined road", "having to shower first thing in the morning", "sitting on the grass talking to a friend" to develop "mental control", and (c) engage in simultaneous feelings awareness and mental control exercises for "meditation proper". Subjects were asked to practise the feelings awareness exercises once or twice a day for no longer than 15 minutes each time.

Group discussions were held at the beginning of each session (except for the first) to answer any queries and clarify any misconceptions subjects might have had. These served as icebreakers and as an opportunity to establish rapport. Here is a session-by-session description of this

program. Table 5.7.1.1 shows a synopsis of this procedure.

Insert Table 5.7.1.1

Session 1

- Personal introduction and statement of the nature, duration and extent of test anxiety for each client. Subjects were asked to talk about other social or evaluative situations which aroused anxiety.
- An explanation of rationale and course of treatment was given.
- Instructed body awareness training. This was introduced as a most important procedure for this treatment as it was a training exercise to promote awareness of the "feelings side" of anxieties. The therapist's task was to direct subject's attention to various parts of their body (hands, arms, head, face, neck) and point out body sensations and feelings (e.g., warm, comfortable. full) associated with that part of the body. Subjects were asked to get as comfortable as possible, to close their eyes, and to take a few minutes to "clear your minds of all distractions and lingering thoughts" before beginning to focus on awareness of feelings and sensations.
- A group discussion of the session followed and the homework was assigned.

Session 2

- The homework assignment was discussed and any misconception of treatment corrected.
-A list of cognitive control scenes was constructed through

TABLE 5.7.1.1

Synopsis of the Placebo procedure

Session 1:

(a) Personal introduction and state of nature, duration and extent of test anxiety for each client (10 min)
(b) Explanation of rationale and course of treatment (5 min)
(c) Instructed body awareness training (30 min)
(d) Group discussion and homework assignment (5 min)

Session 2:

(a) Discussion of homework assignment and correction of any misconceptions of treatment (10 min)
(b) Construction of a list of cognitive control scenes through group discussion (15 min)
(c) Instructed body awareness training (20 min)
(d) Group discussion of session and homework assignment (5 min)

Sessions 3 to 7:

(a) Group discussion of problem areas including changes in cognitive control scenes (10 min)
(b) Instructed body awareness and simultaneous cognitive control exercises (30-35 min)
(c) Group discussion of session and homework assignment (5 min)

---

group discussion. Some of the items were chosen from a list provided and others were constructed by the group. In addition to those mentioned above there were nine others (12 in total), "putting out milk bottles at night", "cleaning your shoes", walking to bed after switching the lights off", etc.

- Instructed body awareness of session and homework assignment.

Sessions 3 to 7

- Group discussions were held into problem areas, misconceptions were clarified and changes or additions made to the list of cognitive control scenes.

- It was first explained (and subjects were consequently reminded) that the simultaneous activities of being aware of body feelings and concentrating on thoughts and cognitions is the only necessary activity for resolving the irrational nature of anxieties such as test anxiety.

- Instructed body awareness and simultaneous cognitive control exercises. After the instructed body awareness phase (which was the same as for session 1) subjects were told that the "thought control exercise" was to begin. First, subjects were instructed to imagine the cognitive scene chosen, in "as vivid details as possible" (the therapist would also provide suggestive details to help create a mental picture of the scene). Then the therapist would proceed to have subjects concentrate on body awareness (e.g., "the sensations in your right hand") and the cognitive scene(e.g., "Imagine you are riding a bicycle along a tree lined road....watch your knees go up and down"). Three cognitive control exercises using three different cognitive scenes were carried out in each of

these sessions.

- A group discussion followed and homework was then assigned. (During the last session a discussion on the program as a whole was also encouraged. No homework assignment was given then).

CHAPTER VI

RESULTS

Unless otherwise specified, all analyses are carried out on the four treatment groups; Study and Exams Skills (group 1), Placebo (group 2), Holroyd's Cognitive-Attentional Training (group 3), and the Cognitive-Attentional Training plus Study and Exams Skills Package (group 4). In those few analyses where the control group is included it is designated as group 5.

In this chapter follow-up is intended to mean the 4-week follow-up, while follow-up 12 is intended to mean the 12-month follow-up.

I shall begin by looking at any differences between groups before treatment (section 6.1). Although administered at follow-up the Ethnocentrism-New Zealand (E-NZ) scale, the Einstellung Water Jug (EWJ) test and the Ray Achievement Motivation (RAM) scale are considered in this section because, as previously pointed out (section 3.6), they are used as covariates in the various analyses and because it is believed that the treatments employed in this study would not affect the construct they were designed to measure; namely, rigidity and achievement motivation.

The Trait anxiety (A-Trait) scale which had been administered at the same time as the E-NZ, the EWJ, and the RAM with similar intentions in mind, was not be treated as a covariate and therefore will not be considered in section 6.1.

This decision was based on the fact that there is clear evidence that general anxiety was affected by treatment.

This is apparent from two points. (i) The A-Trait correlated .57 ($p<.001$) with the cognitive (CSAQ.C) and .26 ($p<.06$) with the somatic scale (CSAQ.S) of the Cognitive Somatic Anxiety Questinnaire (CSAQ) which was administered at pre-test. These correlations are considerably different from those obtained by Schwartz et al. (1978) between the A-Trait and the CSAQ scales (.67 ($p<.001$) and .40 ($p<.001$) respectively for the CSAQ.C and the CSAQ.S). (ii) While the size and significance of the correlations between the two scales of the CSAQ and the three test anxiety scales diminished steadily and considerably from pre-test to post-test and follow-up, the reverse was true for the A-Trait scale (c.f. Table 6.9.2.1). (The effects of treatment on trait anxiety will be reported on section 6.4).

Following the analyses of the preliminary data, I will report the results pertaining to social validity (section 6.2).

In the next four sections of this chapter I will report the results of analyses concerning the effects of treatment on the following constructs: test anxiety, general anxiety, study skills, and academic performance (6.3, 6.4, 6.5, and 6.6 respectively).

In the seventh section (6.7) I will report the results relating to the effects of treatment on the ability measures employed, the verbal and numerical scales of the SCATT. The next three sections relate to correlations between test anxiety measures and performance measures in the various ranges of ability (6.8), the validation of the TASEQ (6.9),

and the effects of treatment on subjects self-efficacy (6.10). In section 6.11 results relating to the follow-up 12 are reported.

To explore more accurately the long term effects of the various treatments, analyses have been carried out (section 6.12) considering the pre-treatment and follow-up 12 test anxiety scores. It was reasoned that, during the period shortly after treatment, test anxiety levels, now well below original levels, would be in a state of flux which would settle by the time follow-up 12 was carried out. In other words, the rationale behind excluding post-test and follow-up test anxiety is that they would bring error into analyses aimed at establishing the long term effects of treatment on test anxiety. Similar analyses would have been carried out on FGA but because very few subjects who were enrolled in 1982 were still enrolled in 1984, these were not warranted.

Finally in section 6.13 I will report the correlations between the covariates here investigated and the test anxiety measures as well as the covariates intercorrelations.

6.1 Preliminaries

Table 6.1.1.1 shows the number of subjects (and percentage for that group) taking courses in the various disciplines. No systematic group difference is apparent on Table 6.1.1.1

Insert Table 6.1.1.1 here

TABLE 6.1.1.1

Number(%) of subjects taking courses in the various disciplines for each group.

Courses in	Group 1 (n=9)	Group 2 (n=8)	Group 3 (n=11)	Group 4 (n=9)	Group 5 (n=22)
Acccountancy	1(11%)				
American Studies			1(9%)	1(11%)	
Art History	1(11%)				
Biology	1(11%)	2(25%)	1(9%)		3(14%)
Business Admin.		1(13%)			
Chemistry	1(11%)	1(13%)			1(5%)
Economics	2(22%)	1(9%)			2(9%)
Education	3(33%)	5(63%)	6(55%)	6(67%)	18(82%)
English	3(33%)	1(13%)	4(36%)	4(44%)	1(5%)
Geology		1(13%)			
Geography			1(9%)		1(5%)
German	1(11%)			4(11%)	
French		1(13%)		1(11%)	
History	1(11%)		1(9%)	1(11%)	3(14%)
Japanese		1(13%)	1(9%)		
Law	1(11%)				
Mathematics			1(9%)		2(9%)
Maori Studies					2(9%)
Music	1(11%)			1(11%)	3(14%)
Philosophy	1(11%)				1(5%)
Physics	1(11%)				
Political Science		2(25%)	3(27%)	1(11%)	1(5%)
Psychology	8(89%)	5(63%)	7(64%)	5(56%)	12(55%)
Religious Studies			1(9%)		2(9%)
Russian				1(11%)	
Sociology		4(50%)	4(36%)	4(44%)	10(45%)
Statistics	3(33%)				
Zoology	1(11%)				

6.1.2 Pre-treatment test anxiety levels

Table 6.1.2.1 shows pre-treatment group means and standard deviations.

Insert Table 6.1.2.1 here

A MANOVA was carried out on the three test anxiety measures here employed to detect any between-group differences at pre-test. As we can see from Table 6.1.2.2 there were no overall significant between-group test anxiety differences. But let's see if there were any between-group differences for each particular measure.

---

insert Table 6.1.2.2 here

Insert Table 6.1.2.3 here

Table 6.1.2.3 shows that even when the three test anxiety measures are considered individually there were no significant between-group differences at pre-test.

As pre-treatment AATD-F levels of the control group were available (mean = 15.09; SD = 5.71; n = 22) an ANOVA was carried out on this measure, including this group as well.

TABLE 6.1.2.1

Test anxiety measures pre-treatment group means(standard deviations)

Group	n	AATD-F	ITA	TASEQ
1	9	16.67(5.83)	44.56(9.96)	962.22(225.76)
2	8	16.13(6.29)	43.63(10.03)	820.00(341.72)
3	11	18.73(7.50)	46.45(12.14)	860.00(318.59)
4	9	16.22(5.80)	42.22(8.63)	753.33(284.52)

TABLE 6.1.2.2

MANOVA on pre-treatment test anxiety measures

Effect	S	M	N	Value	Approx. F	Sig. of F
Group	3	-1/2	14 1/2	.209	.822	.597

TABLE 6.1.2.3

ANOVAs on individual pre-treatment test anxiety measures

Scale	Effect	D.F.	MS	F	Sig. of F
AATD-F	Group	3	15	.361	.782
ITA	Group	3	31	.290	.832
TASEQ	Group	3	68580	.783	.512

Again no between-group difference was apparent, $F(4, 54) = 0.643$, $p < .634$.

6.1.3 Academic performance measures

Two separate ANOVAs (but no MANOVA) were carried out on the two academic performance measures: Tests Exams Average (TEA) and Final Grades Average (FGA). This was done because a sizeable proportion of subjects were not enrolled at university the year before treatment and therefore no pre-treatment FGA was available for them - a MANOVA on TEA and FGA would exclude all these subjects and their contribution to any pre-treatment group differences on pre-test TEA levels would be lost. Moreover, for reasons explained in section 4.3.3 much pre-treatment data on TEA was missing. Finally, pre-treatment FGA but not TEA levels were available for the control group.

Table 6.1.3.1 shows pre-treatment group means and standard deviations. Note how a sizeable number of subjects were not enrolled at university the previous year.

Insert Table 6.1.3.1 here

Table 6.1.3.2 shows the results of the two ANOVAs carried out on FGA and TEA pre-treatment levels respectively.

TABLE 6.1.3.1

Pre-treatment academic performance measures group means

(standard deviations)

F G A			T E A		
Group	n	mean	Group	n	mean
1	5	64.20 (11.54)	1	7	59.93 (12.13)
2	5	59.60 (8.05)	2	6	55.28 (9.28)
3	7	54.71 (8.20)	3	5	57.04 (1.73)
4	6	55.00 (9.40)	4	4	65.42 (1,83)
5	11	57.91 (8.75)	-	-	-

Insert Table 6.1.3.2 here

There is no significant between group difference apparent from Tables 6.1.3.1 and 6.1.3.2.

6.1.4 Scholastic Ability

In order to assess any pre-treatment group differences on scholastic ability measures, a MANOVA and separate ANOVAs were carried out with the verbal (V-SCATT) and the numerical (N-SCATT) scales of the SCATT as dependent measures. Table 6.1.4.1 shows group means and standard deviations on these two scales, Table 6.1.4.2 shows the results of the MANOVA, while Table 6.1.4.3 shows the results of the ANOVAs.

Insert Table 6.1.4.1 here

Insert Table 6.1.4.2 here

Insert Table 6.1.4.3 here

TABLE 6.1.3.2

ANOVAs on pre-treatment FGA and pre-treatment TEA

Measure	Effect	D.F.	MS	F	Sig. of F
FGA	Group	4	84	1.009	.419
TEA	Group	3	91	1.228	.329

TABLE 6.1.4.1

Scholastic ability measures pre-treatment group means (standard deviations)

Group	n	V-SCATT	N-SCATT
1	9	32.11 (6.62)	30.55 (13.11)
2	8	28.88 (5.17)	27.13 (9.42)
3	11	31.36 (7.85)	25.73 (8.58)
4	9	33.44 (8.26)	23.56 (10.53)

TABLE 6.1.4.2

MANOVA on pre-treatment scholastic ability measures

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	2	15	0	6	.144	.852	.535

TABLE 6.1.4.3

ANOVAs on individual pre-treatment scholastic ability measures

Scale	Effect	D.F.	MS	F	Sig. of F
V-SCATT	Group	3	30	.600	.620
N-SCATT	Group	3	78	.713	.551

As we can see from Tables 6.1.4.1, 6.1.4.2 and 6.1.4.3 there were no significant pre-treatment group differences in scholastic ability.

6.1.5 Study and Examination Skills

The three study skills measures; Examination Technique (ET), studying of EXaminations (SE), and Studying Habits (SH), were similarly scrutinized. The respective results are reported in Tables 6.1.5.1, 6.1.5.2 and 6.1.5.3.

Insert Table 6.1.5.1 here

Insert Table 6.1.5.2 here

Insert Table 6.1.5.3 here

From the results reported in this section we can see that there was no significance between group difference at pre-test on the three study skills measures.

TABLE 6.1.5.1

Study skills measures pre-treatment group means (standard deviations)

Group	n	ET	SE	SH
1	9	31.67(7.02)	32.56(6.82)	115.44(21.36)
2	8	34.38(7.63)	32.00(7.05)	133.13(28.06)
3	11	37.55(5.22)	34.18(5.74)	136.45(36.29)
4	9	32.89(5.60)	33.67(7.48)	135.33(23.52)

TABLE 6.1.5.2

MANOVA of pre-treatment study skills measures

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	3	-1/2	14 1/2	9	.272	1.098	.372

TABLE 6.1.5.3

ANOVAs on individual pre-treatment study skills levels

Measure	Effect	D.F.	MS	F	Sig. of F
ET	Group	3	65	1.629	.202
SE	Group	3	9	.204	.893
SH	Group	3	898	1.108	.360

6.1.6 Personality Variables

MANOVA and separate ANOVAs were carried out on the subject variables - CSAQ.C, CSAQ.S, EWJ, E-NZ, and RAM - to ascertain whether there was any systematic (overall or individual) between-group difference on these measures.

Insert Table 6.1.6.1 here

Insert Table 6.1.6.2 here

Insert Table 6.1.6.3 here

As we can see from the above Tables 6.1.6.1, 6.1.6.2 and 6.1.6.3 there was no apparent significant between-group differences for the variables here investigated.

6.2 Social Validity

In this section I am going to report the results from the First Impressions of Treatment (FIT) questionnaire and those from the Participant Reactions Questionnaire (PRQ).

TABLE 6.1.6.1

Personality variables group means (standard deviations)

Group	n	CSAQ.C	CSAQ.S	EWJ	E-NZ	RAM
1	9	19.11(4.01)	17.11(4.78)	1.22(2.16)	-43.56(17.65)	34.67(3.04)
2	8	17.75(3.85)	18.25(3.58)	1.88(2.59)	-48.50(15.97)	29.88(5.74)
3	11	19.50(4.88)	18.40(5.46)	1.30(2.06)	-51.50(13.75)	32.30(5.50)
4	9	21.00(5.66)	20.33(4.46)	1.33(1.97)	-57.67(6.12)	34.50(4.51)

TABLE 6.1.6.2

MANOVA on personality variables measures

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	3	1/2	11 1/2	12	.403	.839	.633

TABLE 6.1.6.3

ANOVAs on individual personality variable measures

Measure	Effect	D.F.	MS	F	Sig. of F
CSAQ.C	Group	3	12	.592	.625
CSAQ.S	Group	3	13	.567	.641
EWJ	Group	3	1	.150	.929
E-NZ	Group	3	255	1.199	.328
RAM	Group	3	40	1.704	.188

6.2.1 Following the first session ....

As the FIT contains only two questions from which data were collected; expectancy, and predisposition for recommending participation to a similar treatment to a friend (FRIEND), they were entered as separate variables in a MANOVA. Separate ANOVAs were carried out on each of these two variables. The results of these analyses are reported on Table 6.2.1.2 and Table 6.2.1.3 respectively. But first, group means and standard deviations are reported (see Table 6.2.1.1).

Insert Table 6.2.1.1 here

Insert Table 6.2.1.2 here

Insert Table 6.2.1.3 here

Although they only approached significance the individual ANOVA clearly showed that it was Expectancy that was contributing strongly to the overall differences, itself standing at significantly different group levels, group 2 being significantly below ($t(1,37) = -2.76$ $p < .009$) and group 4 being almost significantly above ($t(1,37) = 1.98$, $p < .056$)

TABLE 6.2.1.1

Expectancy and FRIEND group means and standard deviations

Group	n	Expectancy	FRIEND
1	9	12.78(6.18)	1.78(.67)
2	8	8.13(3.72)	2.00(0.00)
3	11	13.64(5.95)	1.73(.47)
4	9	15.56(4.64)	1.89(.33)

TABLE 6.2.1.2

MANOVA on Expectancy and FRIEND

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	2	0	15	6	.289	1.862	.101

TABLE 6.2.1.3

ANOVAs on Expectancy and FRIEND

Variable	Effect	D.F.	MS	F	Sig. of F
Expectancy	Group	3	84	2.966	.046
FRIEND	Group	3	0.1	.665	.597

the other three. Group 1 and group 3 were not significantly different from the rest (t (1,37) = 0.17, p< .869; and t(1,37) = .78, p< .443, respectively).

6.2.2 Following treatment ....

A repeated measures ANOVA was carried out on the data collected from the PRQ which will be reported first. A repeated measures MANOVA was then carried out on each of the items in the PRQ except the last, which was identical to FRIEND, thereby providing a pre-test, post-test, and follow-up measurement on this variable.

The PRQ yielded alpha coefficients of .78 at post-test and .89 at follow-up, thereby guaranteeing its reliability and its legitimacy as a research measure.

Tables 6.2.2.1 and 6.2.2.2 show data and analyses concerning the PRQ.

Insert Table 6.2.2.1 here

Insert Table 6.2.2.2 here

TABLE 6.2.2.1

PRQ group means (standard deviations)

at post-test and follow-up

Group	n	post-test	follow-up
1	9	27.11(2.85)	26.33(2.96)
2	8	24.88(4.42)	26.50(4.28)
3	10	27.30(3.37)	26.10(5.40)
4	7	25.29(6.55)	23.00(9.20)

TABLE 6.2.2.2

Repeated measures ANOVA on PRQ

Effect	D.F.	MS	F	Sig. of F
Group	3	23	.614	.611
Time	1	7	.521	.476
Group by Time	3	11	.772	.519

As we can see from Tables 6.2.2.1 and 6.2.2.2 there was no significant effect on the PRQ. But was there any significant effect when we consider each of the PRQ items individually ?

As can be seen from Appendix J, the first item asked subjects to rate the degree to which they felt the treatment had been effective (SUBJ.EVAL.), while the second item probed on the effectiveness of the therapist (THER.EVAL.). The third item probed the therapist's empathy towards them (THER.EMP.) and the fourth on his expertise (THER.EXP.) The fifth item probed the perceived authenticity of treatment (AUTHENTICITY), and the sixth item was the same as FRIEND on the FIT questionnaire thereby allowing a repeated measures ANOVA with three levels instead of two.

Tables 6.2.2.3, 6.2.2.4, and 6.2.2.5 pertain to the results of the PRQ items considered individually.

Insert Table 6.2.2.3 here

---

insert Table 6.2.2.4 here

---

insert Table 6.2.2.5 here

As can be seen from the above three Tables there was no significant effect when these PRQ items were considered

TABLE 6.2.2.3

Post-test and Follow-up group means (standard deviations) on PRQ items (1-5) responses

Item	Time	Group 1	Group 2	Group 3	Group 4
(1) SUBJ. EVAL.	Post	4.78(.67)	4.25(1.04)	4.70(.95)	5.14(.90)
	Follow-up	4.78(.44)	4.75(1.04)	4.80(1.23)	4.57(1.72)
(2) THER. EVAL.	Post	4.89(.60)	4.63(1.41)	4.60(.84)	4.86(.90)
	Follow-up	4.89(.78)	4.88(.99)	5.60(1.17)	4.43(1.51)
(3) THER. EMP.	Post	4.89(1.45)	4.75(.89)	5.70(.82)	4.71(1.60)
	Follow-up	4.44(1.01)	5.25(.71)	5.00(2.00)	4.14(1.77)
(4) THER . EXP.	Post	5.44(.73)	5.25(.89)	5.50(.71)	4.57(1.27)
	Follow-up	5.44(.88)	5.38(.74)	4.90(1.10)	4.86(1.46)
(5) AUTHEN- TICITY	Post	5.11(1.45)	4.25(1.67)	4.80(.92)	4.57(1.27)
	Follow-up	5.22(1.39)	4.63(1.30)	5.00(1.15)	4.14(1.57)

TABLE 6.2.2.4

Repeated measures MANOVA on PRQ items (1-5) responses

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	3	1/2	12	15	.587	1.363	.185
Time	1	1 1/2	12	5	.088	.499	.774
Group by Time	3	1/2	12	15	.392	.843	.629

TABLE 6.2.2.5

Repeated measures ANOVAs on individual PRQ items (1-5)

Item	Effect	D.F.	MS	F	Sig. of F
(1)					
SUBJ.	Group	3	.4	.242	.866
EVAL.	Time	1	.01	.024	.878
	Group by Time	3	.7	1.190	.330
(2)					
THER.	Group	3	.3	.203	.893
EVAL.	Time	1	.01	.020	.887
	Group by Time	3	.3	.406	.749
(3)					
THER.	Group	3	3.8	1.006	.404
EMP.	Time	1	1.8	1.905	.178
	Group by Time	3	1.2	1.321	.286
(4)					
THER.	Group	3	1.5	1.329	.284
EXP.	Time	1	.1	.170	.683
	Group by Time	3	.7	.863	.471
(5)					
AUTHEN-	Group	3	2.4	.987	.412
TICITY	Time	1	.1	.117	.735
	Group by Time	3	.4	.390	.761

simultaneously nor when they are examined individually.

Tables 6.2.2.6 and 6.2.2.7 show the results of FRIEND.

Insert Table 6.2.2.6 here

Insert Table 6.2.2.7 here

As can be seen from the above two tables the group effect approached significance. We shall therefore look at the orthogonal contrast. Group 4 appears to be significantly below the others ($t(1,36) = -2.48$, $p < .020$). None of the other orthogonal contrasts reaches significance ($t(1,36) = 1.06$, $p < .301$; the $t(1,36) = .22$, $p < .824$, $t(1,36) = 1.32$, $p < .199$, respectively for groups 1, 2, and 3).

Table 6.2.2.6 shows that group 4 decreased steadily from pre-test to post-test and follow-up unlike the other three groups. Let's see if the orthogonal contrasts show any significant group by time comparisons. Group 4 appears to be significantly below the others ($t(1,32) = 2.55$, $p < .017$) while none of the other three groups produces a significant orthogonal contrast ($t(1,32) = -1.54$, $p < .135$; $t(1,32) = .37$, $p < .717$; $t(1,32) = -1.48$, $p < .152$; respectively for groups 1, 2, and 3).

Note how group 4 standard deviations at post-test and follow-up were much larger than the other three groups.

TABLE 6.2.2.6

FRIEND Group means (standard deviations) at pre-test, post-test and follow-up

Group	n	Pre-test	Post-test	Follow-up
1	7	1.71(.76)	2.00(0.00)	2.00(0.00)
2	8	2.00(0.00)	1.75(.46)	1.63(.74)
3	9	1.78(.44)	2.00(0.00)	2.00(0.00)
4	7	2.00(0.00)	1.43(1.51)	.86(1.95)

TABLE 6.2.2.7

Repeated measures ANOVA on FRIEND

(Note: 3 levels)

Effect	D.F.	MS	F	Sig. of F
Group	3	1.2	2.232	.107
Time	2	.4	.690	.506
Group by Time	6	.8	.832	.244

6.3 Effects of Treatment and Moderator Variables on Test Anxiety

In this section I will report the results pertaining to the effects of treatment on test anxiety. First, I will report the outcome of a MANOVA on the three test anxiety measures and the results of individual ANOVAs on each of them. Next, I will report the results concerning the moderating effects of the various subjects variables on treatment effects (MANCOVA). Each covariate was tested for its significance, and the significant or near significant covariates entered simultaneously in one analysis to see whether any were redundant. Those that remain significant following this step will be called the selected covariates. Finally each of the covariates that failed to reach significance by themselves are entered in the ANOVA, together with the selected covariates to see if any covariate that individually failed to reach significance enhanced the significance of the selected covariates regression or the size of the multiple R.

The covariates that are found to contribute significantly to the test anxiety measures regression following all the above steps will be called the salient covariates.

6.3.1 Effects of treatment on test anxiety: no moderator variables taken into account

Tables 6.3.1.1, 6.3.1.2, 6.3.1.3, show the results of treatment on test anxiety and Figures 6.3.1.1, 6.3.1.2, and 6.3.1.3 portray them graphically. We would expect scores on the AATD-F and the ITA to decrease following treatment but because of its nature TASEQ scores are expected to increase.

Insert Table 6.3.1.1 here

Insert Figure 6.3.1.1 here

Insert Figure 6.3.1.2 here

Insert Figure 6.3.1.3 here

Insert Table 6.3.1.2 here

TABLE 6.3.1.1

Test anxiety measures pre-test, post-test, and follow-up group means (standard deviations)

Group	n	Time of measurement	AATD-F	ITA	TASEQ
1	9	Pre-test	16.67(5.83)	44.56(9.96)	962.22(225.76)
		Post-test	5.78(9.52)	36.55(8.00)	1291.11(249.52)
		Follow-up	6.00(12.79)	37.00(11.70)	1207.78(282.03)
2	8	Pre-test	16.13(6.29)	43.63(10.03)	820.00(341.72)
		Post-test	7.50(7.03)	41.75(8.60)	1053.75(394.02)
		Follow-up	2.25(13.04)	33.13(7.99)	1265.00(325.01)
3	11	Pre-test	18.73(7.50)	46.45(12.14)	860.00(318.59)
		Post-test	8.55(8.39)	44.55(12.98)	1436.36(324.32)
		Follow-up	7.09(6.63)	37.45(9.77)	1363.64(425.05)
4	9	Pre-test	16.22(5.80)	42.22(8.63)	753.33(284.52)
		Post-test	3.44(12.79)	35.11(11.15)	1345.56(528.59)
		Follow-up	-1.22(13.11)	28.22(9.18)	1537.78(364.99)

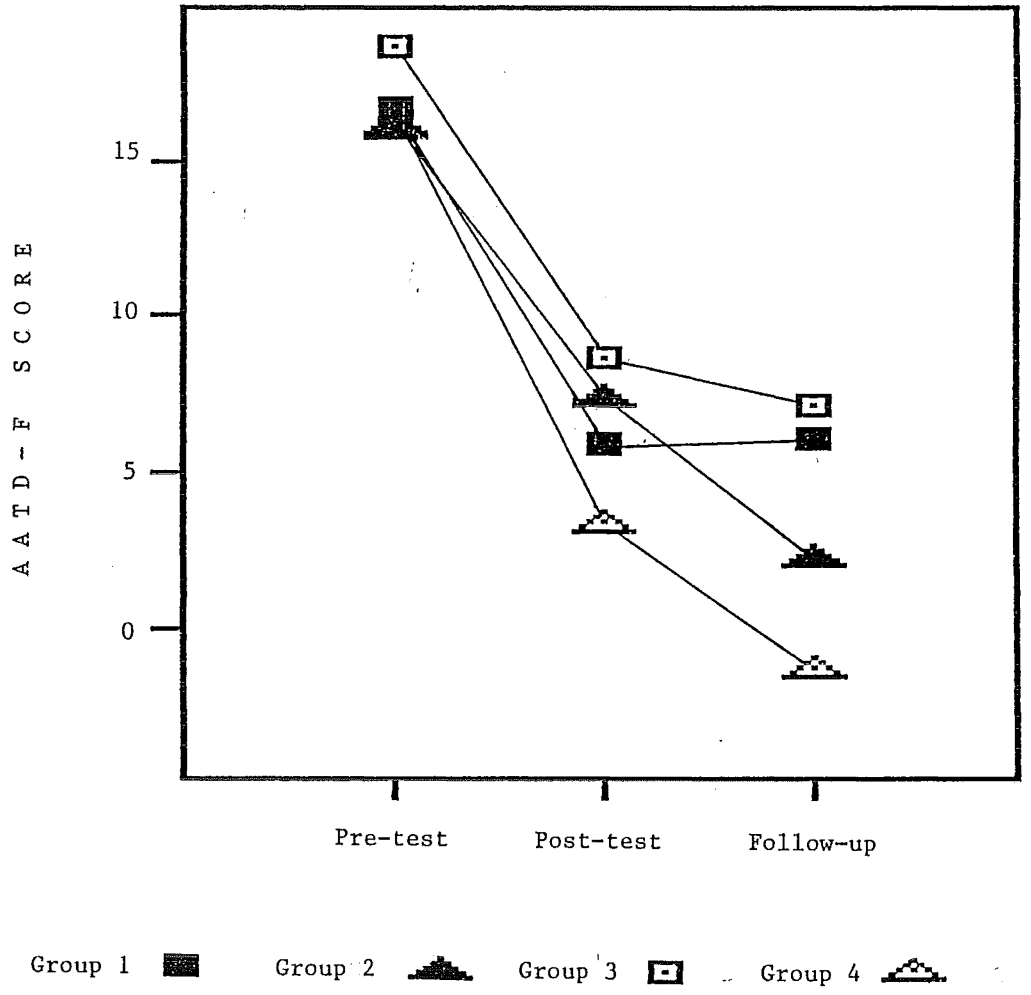
FIGURE CAPTIONS

Figure 6.3.1.1: AATD-F Group Means at Pre-test, Post-test, and Follow-up.

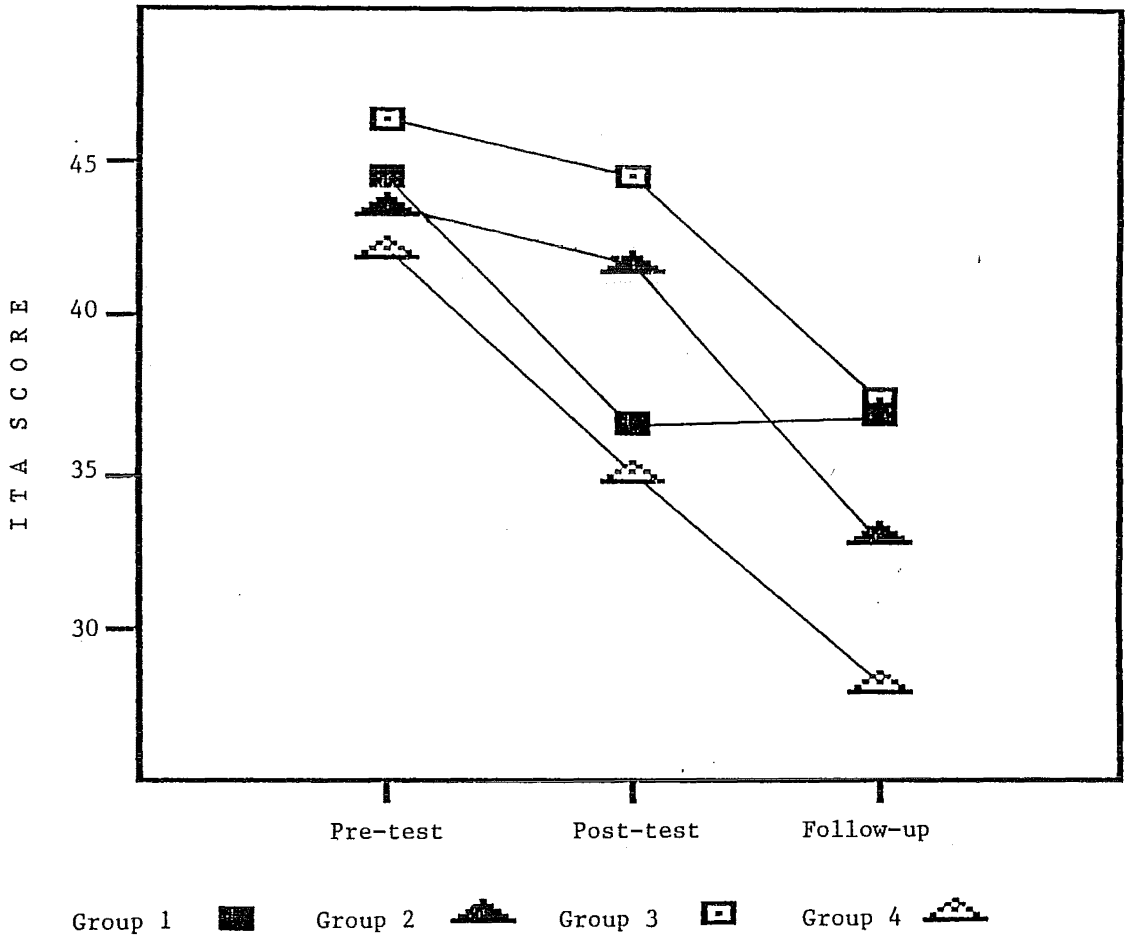
Figure 6.3.1.2: ITA Group Means at Pre-test, Post-test, and Follow-up.

Figure 6.3.1.3: TASEQ Group Means at Pre-test, Post-test, and Follow-up.

F I G U R E 6.3.1.1


A A T D - F S C O R E
15
10
5
0
Pre-test
Post-test
Follow-up
Group 1
Group 2
Group 3
Group 4

F I G U R E 6.3.1.2


ITA SCORE
45
40
35
30
Pre-test
Post-test
Follow-up
Group 1
Group 2
Group 3
Group 4

F I G U R E 6.3.1.3.

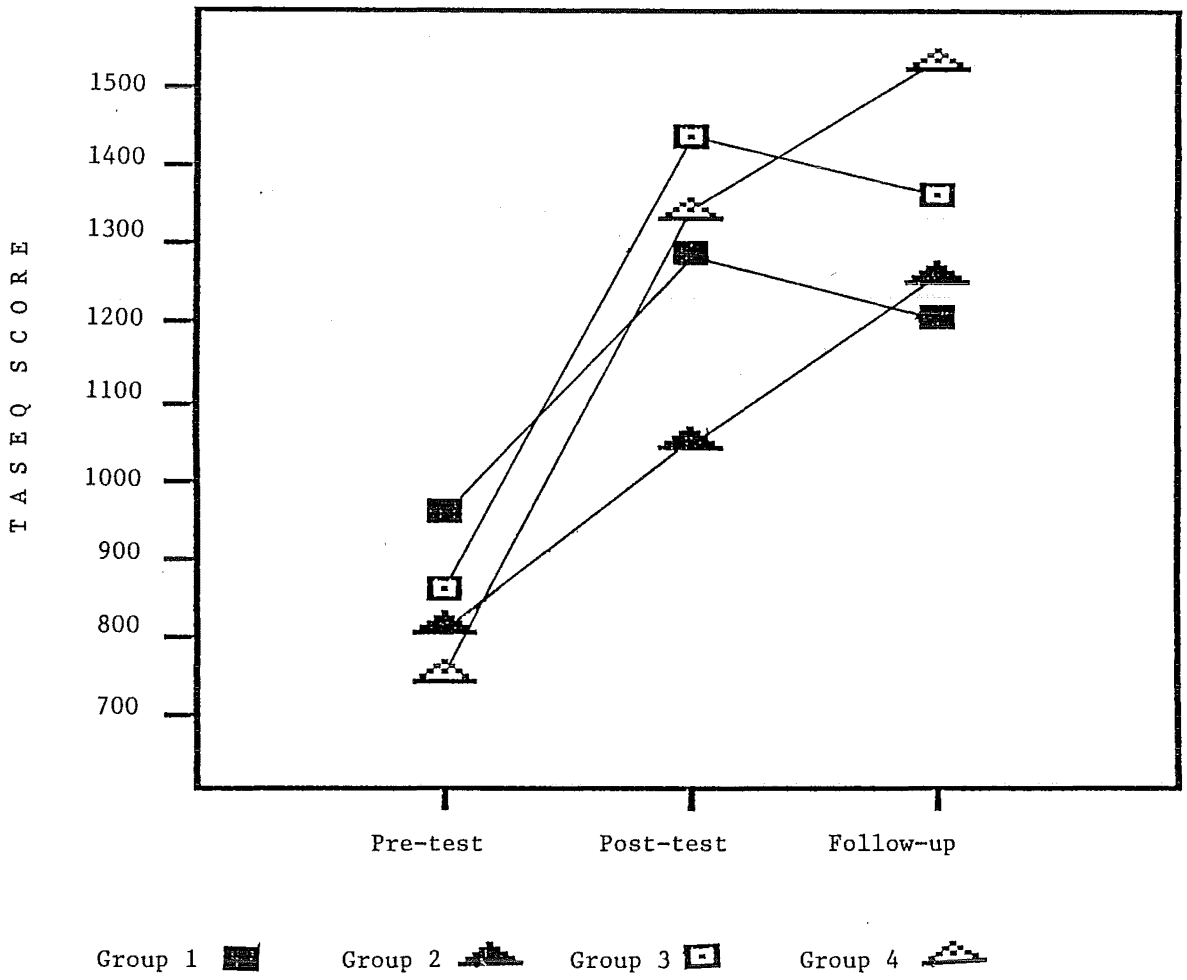

T A S E Q S C O R E
1500
1400
1300
1200
1100
1000
900
800
700
Pre-test
Post-test
Follow-up
Group 1
Group 2
Group 3
Group 4

TABLE 6.3.1.2

Repeated measures MANOVA on test anxiety measures.

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	3	-1/2	14 1/2	9	.250	1.001	.444
Group	1	2	13	6	.779	16.422	.001
Group by Time	3	1	13	18	.566	1.162	.309

Insert Table 6.3.1.3 here

As can be seen from the above three tables there was no significant effect for group, be it at a multivariate or a univariate level. There was a highly significant effect for time both at multivariate and at univariate levels. The group by time effect did not approach significance on the multivariate analysis nor on two of the ANOVAs (for AATD-F and for ITA), nonetheless as Table 6.3.1.3 shows there was a virtually significant group by time effect on the TASEQ.

An examination of the orthogonal contrasts shows group 4 to have increased on this measure at a significantly higher rate when compared with the other three groups ($t(1,33) = -2.40$, $p < .022$) while group 1 increased at a significantly lower rate than the other three ($t(1,33) = 2.06$, $p < .047$). Neither group 2 nor group 3 orthogonal contrasts approached significance ($t(1,33) = .39$, $p < .696$, and $t(1,33) = -.08$, $p < .937$, respectively).

6.3.2 Did individual groups improve on test anxiety?

Tables 6.3.2.1 and 6.3.2.2 show respectively the results of MANOVAs and ANOVAs carried out on each group taken individually.

TABLE 6.3.1.3

Repeated measures ANOVAs on test anxiety measures

Measure	Effect	D.F.	MS	F	Sig. of F
AATD-F	Group	3	142	.720	.547
	Time	2	1833	53.802	.0001
	Group by time	6	28	.827	.553
ITA	Group	3	288	1.352	.275
	Time	2	963	18.670	.001
	Group by time	6	53	1.024	.418
TASEQ	Group	3	167293	.767	.521
	Time	2	2762179	37.862	.0001
	Group by time	6	161871	2.219	.052

Insert Table 6.3.2.1 here

Insert Table 6.3.2.2 here

We can see that, with one exception (group 1 on ITA ANOVA), effects for time tended to be highly significant.

6.3.3 Looking for significant covariates

5.3.3.1 shows each of the investigated covariates regression multivariate test of significance. CSAQ-Trait is the sum of the A-Trait variance explained by the two CSAQ sclaes (CSAQ.C x.67^2 + CSAQ.S x .40^2) (c.f. section 4.4.1).

Insert Table 6.3.3.1 here

E-NZ and V-SCATT were the only significant covariates at the multivariate level while CSAQ.C, CSAQ-Trait, N-SCATT, and SH approached significance. A new MANCOVA was carried out with E-NZ, V-SCATT, CSAQ-Trait, N-SCATT, and SH (the selected covariates) entered simultaneously. CSAQ.C was left out because of its obvious similarity to CSAQ-Trait which was

TABLE 6.3.2.1

Repeated measures MANOVAs on test anxiety measures for each of the four treatment groups taken individually

Group	Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
1	Time	2	0	6	6	.676	2.552	.041
2	Time	2	0	5	6	.914	3.644	.009
3	Time	2	0	8	6	.909	5.275	.001
4	Time	2	0	6	6	.917	4.230	.003

TABLE 6.3.2.2

Repeated measures ANOVAs on the three test anxiety measures for each of the four treatment groups taken individually

Measure	Group	Effect	D.F.	MS	F	Sig. of F
AATD-F	1	Time	2	34	10.237	.001
	2	Time	2	32	12.453	.001
	3	Time	2	25	18.013	.001
	4	Time	2	48	15.228	.001
ITA	1	Time	2	74	2.472	.116
	2	Time	2	41	6.188	.012
	3	Time	2	65	3.792	.040
	4	Time	2	22	19.938	.001
TASEQ	1	Time	2	30127	8.733	.003
	2	Time	2	63507	6.242	.012
	3	Time	2	72489	14.951	.001
	4	Time	2	124632	12.072	.001

TABLE 6.3.3.1

Covariates regressions multivariates tests of significance

Covariate	S	M	N	D.F.	Value	Approx. F	Sig. of F
E-NZ	1	1/2	13 1/2	3	.298	4.110	.015
EWJ	1	1/2	12	3	.204	2.226	.109
CSAQ.C	1	1/2	14	3	.192	2.376	.090
CSAQ.S	1	1/2	14	3	.116	1.317	.287
CSAQ-Trait	1	1/2	14	3	.197	2.451	.083
ACHMOT	1	1/2	13 1/2	3	.091	.971	.420
V-SCATT	1	1/2	14	3	.234	3.053	.044
N-SCATT	1	1/2	14	3	.205	2.574	.072
SH	1	1/2	14	3	.203	2.550	.074
SE	1	1/2	14	3	.013	.127	.943
ET	1	1/2	14	3	.037	.383	.766
Expectancy	1	1/2	14	3	.141	1.639	.201

more closely approaching significance. Tables 6.3.3.2 and 6.3.3.3 show respectively the multivariate and the univariate tests of significance of the above analysis. On the other hand Table 6.3.3.4 examines each of the selected covariates contribution and significance to each of the multiple regressions with the test anxiety measures as dependent variables individually. Again because of the nature of the TASEQ a significant covariate was expected to yield Beta weights of opposite sign when compared with the AATD-F and the ITA.

Insert Table 6.3.3.2 here

Insert Table 6.3.3.3 here

Insert Table 6.3.3.4 here

We can see that N-SCATT and SH did not contribute significantly to any of the three multiple regressions.

The same MANCOVA as above was carried out but with only the actual and not spurious moderator variables, that is E-NZ, CSAQ-Trait, and V-SCATT. Table 6.3.3.5 reports the results of this analysis and it reveals that the size of these multiple regressions was virtually unchanged but their

TABLE 6.3.3.2

Significant covariates
(E-NZ, V-SCATT, CSAQ-Trait, N-SCATT, SH)
regression mutlti-variate test of significance

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
(Regression)	3	1/2	11 1/2	15	.903	2.326	.008

TABLE 6.3.3.3

Significant covariates (E-NZ, V-SCATT, CSAQ-Trait, N-SCATT, SH) multiple regressions, univariate tests of significance:
5,27 degrees of freedom.

Dep. Variable	Sq. Mul. R	Mul. R	MS	F	Sig. of F
AATD-F	.420	.648	544	3.904	.009
ITA	.405	.636	142	3.675	.012
TASEQ	.441	.664	632898	4.251	.006

TABLE 6.3.3.4

Analyses of the significant covariates (E-NZ, V-SCATT, CSAQ-Trait, N-SCATT, SH) multiple regressions

Dep. Variable	Covariate	Beta	T-Value	Sig. of F
AATD-F	E-NZ	-.463	-2.732	.011
	CSAQ-Trait	.139	.749	.461
	V-SCATT	-.278	-1.584	.125
	N-SCATT	-.090	- .481	.634
	SH	.004	.023	.982
ITA	E-NZ	-.303	-1.766	.089
	CSAQ-Trait	.399	2.128	.043
	V-SCATT	-.412	-2.315	.028
	N-SCATT	.257	1.362	.184
	SH	.019	.097	.923
TASEQ	E-NZ	.373	2.244	.033
	CSAQ-Trait	-.168	- .925	.363
	V-SCATT	.498	2.886	.008
	N-SCATT	-.195	-1.068	.295
	SH	.120	.646	.524

significance was further enhanced.

Insert Table 6.3.3.5 here

The next step in this investigation was to see if any of those covariates that individually were neither significant nor approaching significance would when entered together with the salient covariates in the MANCOVA significantly contribute to any of the individual multiple regressions (i.e., with any of AATD-F, ITA, or TASEQ as dependent variables) and would appreciably enhance the size of these multiple Rs. Tables 6.3.2.6 through to 6.3.3.10 shows the results for each of the relevant analyses.

Insert Table 6.3.3.6 here

Insert Table 6.3.3.7 here

Insert Table 6.3.3.8 here

Insert Table 6.3.3.9 here

TABLE 6.3.3.5

Selected covariates (E-NZ, CSAQ-Trait, V-SCATT) multiple regressions univariate tests of significance:
3,29 degrees of freedom

Dep. Variable	Sq. Mul. R	Mul. R	MS	F	Sig. of F
AATD-F	.414	.643	894	6.825	.001
ITA	.361	.601	774	5.464	.004
TASEQ	.394	.628	150132	6.283	.002

TABLE 6.3.3.6

Analysis of the three selected covariates plus EWJ multiple regression

Dep. Variable (Mul. R, F, Sig. of F)	Covariate	Beta	T-Value	Sig. of F
AATD-F	E-NZ	-.422	-2.621	.015
(R = .673	CSAQ-TRAIT	.305	1.764	.090
F(4,25)=5.176,	V-SCATT	-.251	-1.623	.117
p< .004)	EWJ	-.251	-1.553	.133
ITA	E-NZ	-.154	- .976	.338
(R = .689,	CSAQ-Trait	.577	3.402	.002
F(4,25)=5.650	V-SCATT	-.136	- .899	.377
p<.002)	EWJ	.007	.043	.966
TASEQ	E-NZ	.187	1.153	.260
(R = .667,	CSAQ-Trait	-.462	-2.652	.014
F(4,25)=4.999,	V-SCATT	.315	2.024	.054
p<.004)	EWJ	.089	.546	.590

TABLE 6.3.3.7

Analysis of the three selected covariates plus SE multiple regression

Dep. Variable (Mul. R, F, Sig. of F)	Covariate	Beta	T-Value	Sig. of T
AATD-F	E-NZ	-.510	-3.196	.003
(R = .644,	CSAQ-Trait	.125	.778	.443
F(4,28)=4.969,	V-SCATT	-.317	-2.164	.039
p<.004)	SE	-.039	- .252	.803
ITA	E-NZ	-.212	-1.272	.214
(R = .601,	CSAQ-Trait	.381	2.266	.031
F(4,28)=3.964,	V-SCATT	-.309	-2.020	.053
p<.011)	SE	-.022	- .134	.895
TASEQ	E-NZ	.302	1.868	.072
(R = .630,	CSAQ-Trait	-.211	-1.290	.208
F(4,28)=4.618,	V-SCATT	.454	3.057	.005
p<.005)	SE	.064	.405	.689

TABLE 6.3.3.8

Analysis of the three selected covariates plus ET multiple regression

Dep. Variable (Mul. R, F, Sig of F)	Covariate	Beta	T-Value	Sig. of T
AATD-F	E-NZ	- .504	-3.225	.003
(R = .644,	CSAQ-Trait	.127	.776	.445
F(4,28)=4.957,	V-SCATT	- .315	-2.156	.040
p<.004)	ET	- .029	-1.88	.852
ITA	E-NZ	- .212	-1.296	.206
(R = .602,	CSAQ-Trait	.378	2.213	.035
F(4,28)=3.968,	V-SCATT	- .308	-2.025	.053
p<.011)	ET	- .028	- .173	.864
TASEQ	E-NZ	.312	1.986	.057
(R = .640,	CSAQ-Trait	- .182	-1.111	.276
F(4,28)=4.857,	V-SCATT	.456	3.112	.004
p<.004)	ET	.134	.863	.396

TABLE 6.3.3.9

Analysis of the three salient covariates
plus Expectancy multiple regression

Dep. Variable (Mul. R, F, Sig. of F)	Covariate	Beta	T-Value	Sig. of T
AATD-F	E-NZ	.485	-3.192	.003
(R = .654,	CSAQ-Trait	.080	.475	.638
F(4,28)=5.231,	V-SCATT	-.294	-2.020	.053
p<.003)	Expectancy	-.135	- .823	.417
ITA	E-NZ	-.198	-1.240	.226
(R = .605,	CSAQ-Trait	.355	2.013	.054
F(4,28)=4.034,	V-SCATT	-.296	-1.931	.064
p<.010)	Expectancy	-.077	- .446	.659
TASEQ	E-NZ	.273	1.758	.090
(R = .635,	CSAQ-Trait	-.182	-1.067	.295
F(4,28)=4.735,	V-SCATT	.432	2.906	.007
p<.005)	Expectancy	.113	.670	.508

Insert Table 6.3.3.10 here

Insert Tables 6.3.3.11 here

As we can see from the tables ACHMOT was the only covariate to satisfy the criteria set above. Therefore it appears that E-NZ, V-SCATT, CSAQ-Trait, and ACHMOT were the subject variables that significantly moderate the effect of test anxiety treatment: The salient covariates.

Table 6.3.3.12 shows the univariate tests of significance for the multiple regression formed by E-NZ, V-SCATT, CSAQ-Trait, and ACHMOT simultaneously entered in the MANCOVA. The multivariate test for this multiple regression can be found in Table 6.3.4.2.

Insert Table 6.3.3.12 here

TABLE 6.3.3.10

Analysis of three selected covariates plus ACHMOT multiple regression

Dep. Variable (Mul. R, F, Sig. of F)	Covariate	Beta	T-Value	Sig. of T
AATD-F	E-NZ	-.501	-3.531	.001
(R = .705,	CSAQ-Trait	.084	.580	.567
F(4,28)=6.909,	V-SCATT	-.355	-2.604	.015
p<.001)	ACHMOT	-.295	-2.147	.041
ITA	E-NZ	-.207	-1.339	.191
(R = .632,	CSAQ-Trait	.351	2.222	.035
F(4,28)=4.663,	V-SCATT	-.336	-2.254	.032
p<.005)	ACHMOT	-.201	-1.343	.190
TASEQ	E-NZ	.285	1.853	.075
(R = .639,	CSAQ-Trait	-.208	-1.323	.196
F(4,28)=4.837,	V-SCATT	.465	3.148	.004
p<.004)	ACHMOT	.124	.835	.411

TABLE 6.3.3.11

Analysis of the three selected covariates plus CSAQ.S multiple regression

Dep. Variable (Mul. R, F, Sig. of F)	Covariate	Beta	T-Value	Sig. of T
AATD-F	E-NZ	-.500	-3.260	.003
(R = .644,	CSAQ-S	.160	.743	.463
F(4,28)=4.952,	V-SCATT	-.307	-2.051	.050
p<.004)	CSAQ-S	.033	- .154	.878
ITA	E-NZ	-.203	-1.269	.215
(R = .603,	CSAQ-S	.339	1.505	.144
F(4,28)=3.994,	V-SCATT	-.318	-2.036	.051
p<.011)	CSAQ.S	.068	.310	.759
TASEQ	E-NZ	.288	1.858	.074
(R = .632,	CSAQ-S	-.310	-1.418	.167
F(4,28)=4.662,	V-SCATT	.429	2.828	.009
p<.005)	CSAQ.S	.112	.521	.606

TABLE 6.3.3.12

E-NZ, V-SCATT, CSAQ-Trait, and ACHMOT regression univariate tests of significance: 4,28 degrees of freedom

Dep. Variable	Sq. Mul. R	Mul. R	MS	F	Sig. of F
AATD-F	.497	.705	116	6.909	.001
ITA	.400	.639	137	4.663	.005
TASEQ	.409	.639	151721	4.837	.004

6.3.4 Effects of treatments on test anxiety controlling for moderator variables

This section is the equivalent of section 6.3.1 except that the moderator variables are controlled for. Table 6.3.4.1 shows the pre-test, post-test, and follow-up adjusted means for the three test anxiety measures while Figures 6.3.4.1, 6.3.4.2, and 6.3.4.3 portray them graphically.

Insert Table 6.3.4.1 here

Insert Figure 6.3.4.1 here

Insert Figure 6.3.4.2 here

Insert Figure 6.3.4.3 here

Table 6.3.4.2 shows the results of the MANCOVA. Note how the multivariate effect for group was nearly significant now but the effects for time and group by time have hardly changed.

TABLE 6.3.4.1

Test anxiety measures pre-test, post-test, and follow-up group means adjusted for the four salient covariates (E-NZ, CSAQ-Trait, V-SCATT, ACHMOT)

Group	n	Time of measurement	AATD-F	ITA	TASEQ
1	9	pre-test	18.86	46.08	899.17
		post-test	9.57	39.21	1197.97
		follow-up	10.72	39.58	1144.09
2	8	pre-test	15.27	44.63	791.17
		post-test	5.35	38.66	1121.19
		follow-up	.87	31.50	1331.69
3	11	pre-test	18.26	46.46	869.53
		post-test	7.71	44.04	1455.43
		follow-up	6.12	37.11	1373.84
4	8	pre-test	15.03	38.15	799.43
		post-test	1.76	32.55	1363.64
		follow-up	-4.33	25.59	1534.68

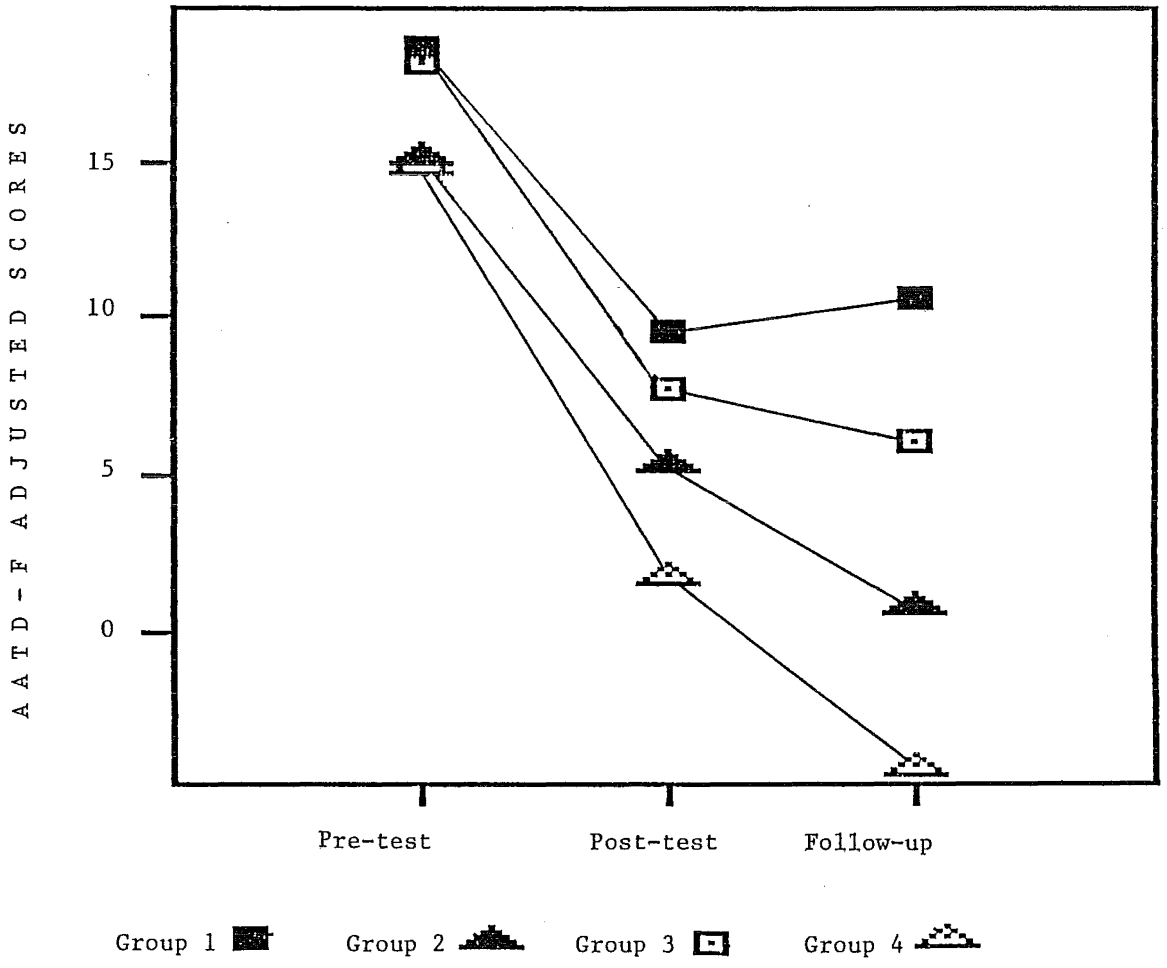
FIGURE CAPTIONS

Figure 6.3.4.1: AATD-F Group Means at Pre-test Post-test, and Follow-up, adjusted for E-NZ, CSAQ-Trait, V-SCATT, ACHMOT.

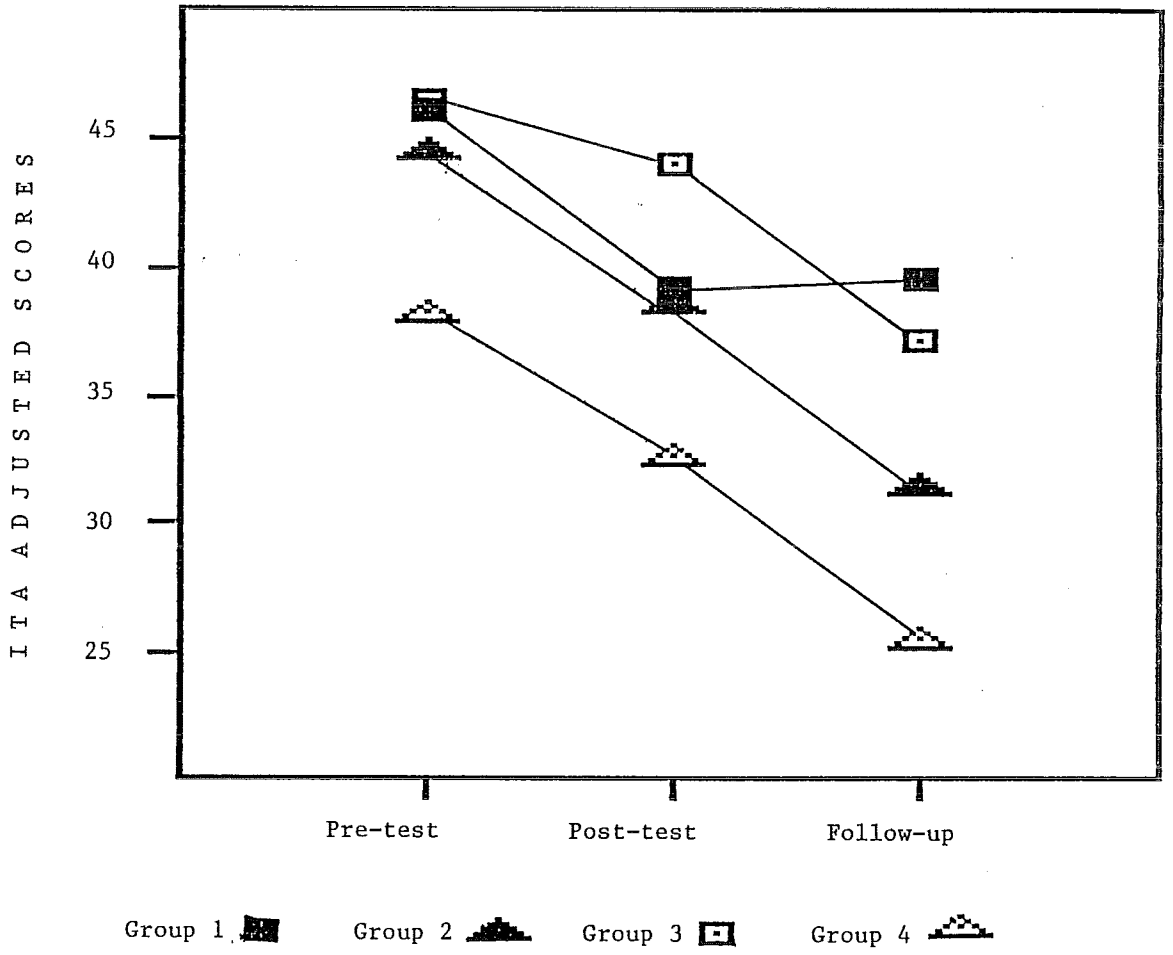
Figure 6.3.4.2 ITA Group means at Pre-test, Post-test, and Follow-up, adjusted for E-NZ, CSAQ-Trait, V-SCATT, ACHMOT.

Figure 6.3.4.3: TASEQ Group Means at Pre-test, Post-test, and Follow-up, adjusted for E-NZ, CSAQ-Trait, V-SCATT, ACHMOT.

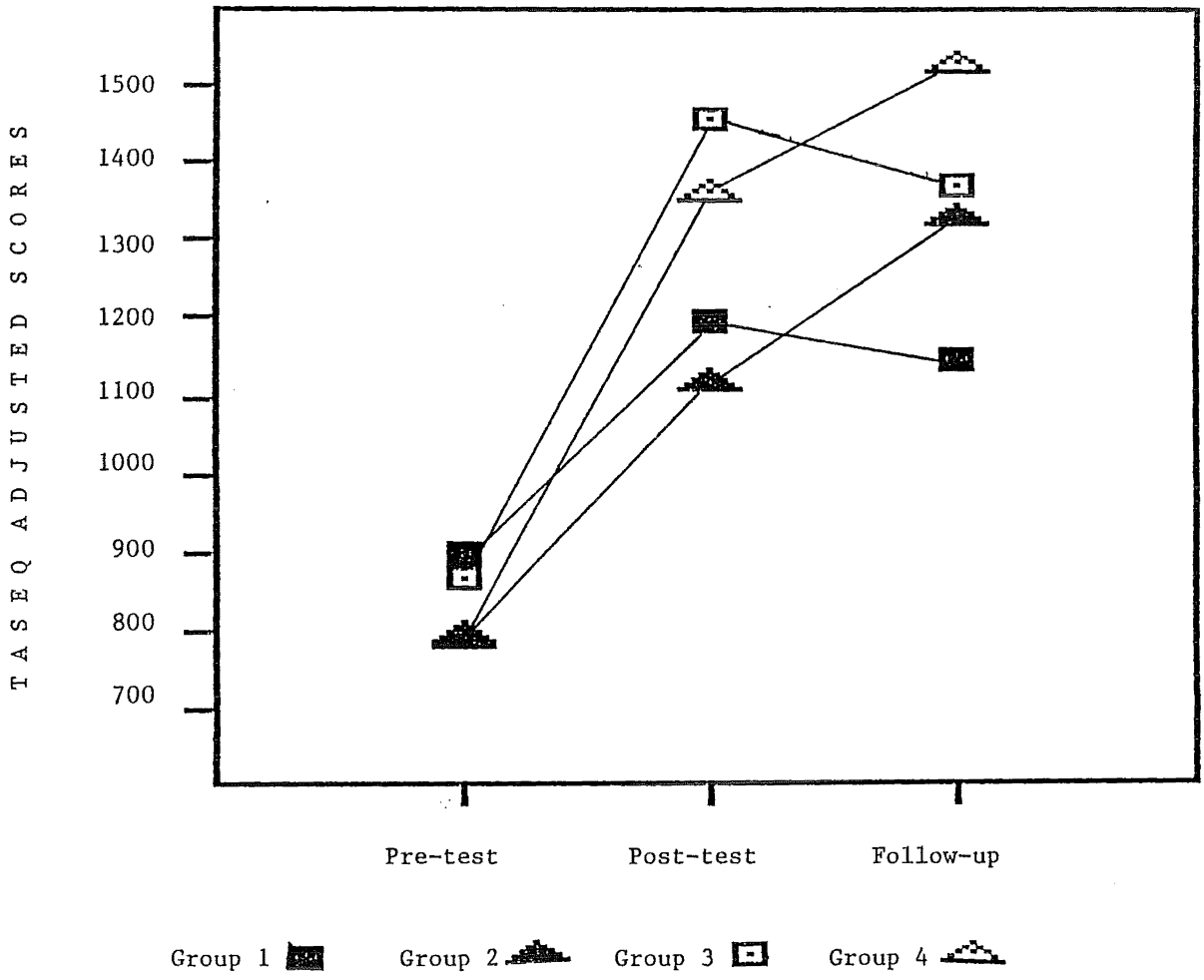
F I G U R E 6.3.4.1.


A A T D - F A D J U S T E D S C O R E S
15
10
5
0
Pre-test
Post-test
Follow-up
Group 1
Group 2
Group 3
Group 4

F I G U R E 6.3.4.2.


I T A A D J U S T E D S C O R E S
45
40
35
30
25
Pre-test
Post-test
Follow-up
Group 1
Group 2
Group 3
Group 4

F I G U R E 6.3.4.3.


T A S E E Q A D J U S T E D S C O R E S
1500
1400
1300
1200
1100
1000
900
800
700
Pre-test
Post-test
Follow-up
Group 1
Group 2
Group 3
Group 4

Insert Table 6.3.4.2 here

Insert Table 6.3.4.3 here

Table 6.3.4.3 shows the results of the individual ANCOVAs. Note how the effect for group now was significant for AATD-F and ITA, while although still not significant, the significance level for TASEQ was now .331 compared to .521 when no covariates were introduced (c.f. Table 6.3.1.3). No appreciable change took place in the time effect which was highly significant anyway. A slight change towards greater significance took place in the group by time effect on all three measures. The TASEQ group by time effect, previously almost significant, was now within significance levels. Table 6.3.3.4 shows the orthogonal contrasts for both the group and the group by time effects.

Wherever there was a significant effect on any of the ANCOVAs there was a corresponding significant contrast showing that group 4 benefited from treatment to a larger extent than the other groups. In addition there were three other significant or near significant contrasts on Table 6.3.4.4. In those cases the group concerned indicates significant or near significant levels of improvement below

TABLE 6.3.4.2

Repeated measures MANCOVA on test anxiety measures
with the salient moderator variables
(E-NZ, V-SCATT, CSAQ-Trait, ACHMOT) as covariates

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
(Regression)	3	0	12	12	.790	2.502	.007
Group	3	-1/2	12	9	.500	1.866	.068
Time	1	2	12 1/2	6	.776	15.575	.001
Group by Time	3	1	12 1/2	18	.581	1.161	.311

TABLE 6.3.4.3

Repeated measures ANCOVAs on test anxiety measures with the four salient covariates (E-NZ, V-SCATT, CSAQ-Trait, ACHMOT)

Measure	Effect	D.F.	MS	F	Sig. of F
AATD-F	(Regression)	4	804	6.909	.001
	Group	3	348	2.990	.048
	Time	2	35	50.818	.0001
	Group by Time	6	29	.817	.561
ITA	(Regression)	4	642	4.663	.005
	Group	3	518	3.759	.022
	Time	2	53	17.689	.001
	Group by Time	6	55	1.030	.414
TASEQ	(Regression)	4	733914	4.837	.004
	Group	3	180751	1.191	.331
	Time	2	73893	36.839	.0001
	Group by Time	6	172594	2.336	.042

the other groups. The groups concerned were group 1 on the group effect for AATD-F and the group by time effect on the TASEQ, and group 3 on the group effect for ITA.

Insert Table 6.3.4.4 here

6.3.5 Worry versus emotionality

In order to test whether the two groups receiving cognitive-attentional training (ie, PK and CATH) showed decreases on the worry scale (W-ITA) greater than on the emotionality scale (E-ITA) of the ITA as Liebert and Morris's (1967) theory of test anxiety would predict (c.f. section 1.3), the following manipulations were carried out. Firstly, two psuedogroups (W and E) were formed of 20 subjects each. The subjects were exactly the same for both pseudogroups (11 from group 3 plus nine from group 4). Secondly, a pseudovariable (Z) was formed: for pseudogroup W, Z was equal to the subject's W-ITA score but for pseudogroup E, Z was equal to the subject's E-ITA score.

Now, an ANOVA on pseudogroups W and E for Z would indicate by means of a group effect or a group by time effect whether subjects receiving CATH had differentially decreased on W-ITA or E-ITA. For instance a larger mean for

TABLE 6.3.4.4

Between groups orthogonal contrasts for test anxiety measures adjusted for the salient moderator variables (E-NZ, V-SCATT, CSAQ-Trait, ACHMOT): 1,36 degrees of freedom

Variable	Group	Effect	T-Value	Sig. of T
AATD-F	1	Group	2.179	.038
		Group by Time	.919	.366
	2	Group	-.769	.449
		Group by Time	.121	.905
	3	Group	1.133	.267
		Group by Time	-.695	.493
	4	Group	-2.215	.035
		Group by Time	1.315	.199
ITA	1	Group	1.341	.191
		Group by Time	-.947	.352
	2	Group	.007	.995
		Group by Time	.035	.972
	3	Group	2.039	.051
		Group by Time	-.536	.596
	4	Group	-2.941	.006
		Group by Time	1.297	.205
TASEQ	1	Group	-1.081	.289
		Group by Time	1.950	.061
	2	Group	-1.000	.325
		Group by Time	.425	.674
	3	Group	1.249	.222
		Group by Time	.013	.990
	4	Group	1.004	.324
		Group by Time	-2.279	.031

pseudogroup W and a significant group or group by time effect on the ANOVA would indicate that CATH decreased worry to a greater extent than emotionality. I wish to remind the reader that both scales of the ITA have eight items each and that for each scale the possible range of scores was 8-40 (c.f. section 4.1.2). Tables 6.3.5.1 and 6.3.5.2 show the relevant means (and standard deviations) and the results of the ANOVA respectively.

Insert Table 6.3.5.1 here

Insert Table 6.3.5.2 here

We can see that there was a significant effect for group and a highly significant effect for time, but no group by time significant effect.

6.3.6 Self-efficacy in not worrying versus self-efficacy in not becoming emotional

The same analysis as in the previous section was carried out on the worry and emotionality scales of the TASEQ.

The worry scale (w-TASEQ) yielded alpha coefficients of .80, .95, and .96 at the pre-test, post-test, and follow up respectively. The corresponding alpha coefficients for the

TABLE 6.3.5.1

Pseudovariable Z means (standard deviations) at pre-test, post-test, and follow-up for pseudogroups W and E

Pseudogroup	n	pre-test	post-test	follow-up
W	20	24.40(6.06)	21.90(7.20)	19.10(6.06)
E	20	20.15(5.52)	18.40(6.58)	14.20(6.41)

TABLE 6.3.5.2

Repeated measures ANOVA on pseudovariable Z for pseudogroups W and E

Effect	D.F.	MS	F	Sig. of F
Group	1	533	6.159	.018
Time	2	323	21.251	.0001
Group by Time	2	5	.323	.725

emotionality scale (E-TASEQ) were .79, .94, and .93.

Executing the same manipulations as for the W-ITA and the E-ITA we obtained two pseudogroups (WW and EE), which of course were composed of the same subjects except that the transformed pseudovariable ZZ was equal to the W-TASEQ for the WW pseudogroup and equal to twice the E-TASEQ for the EE pseudogroup. Because the number of items and possible range of scores was exactly twice as much on the W-TASEQ as on the E-TASEQ (c.f. section 4.2.1), the scores on the latter scale were double to give a comparable index.

Tables 6.3.6.1 and 6.3.6.2 show the relevant means (standard deviations) and the results of the ANOVA respectively.

Insert Table 6.3.6.1 here

Insert Table 6.3.6.2 here

We can see that there was no significant group or group by time effect evident from this analysis.

TABLE 6.3.6.1

Pseudovariable Z Z means (standard deviations) at pre-test, post-test, and follow-up for pseudogroups WW and EE

Pseudogroups	n	pre-test	post-test	follow-up
WW	20	578.50 (199.77)	949.50 (264.66)	1005.50 (233.81)
EE	20	477.00 (226.36)	882.00 (336.29)	873.00 (358.08)

TABLE 6.3.6.2

Repeated measures ANOVA on pseudovariable Z for pseudogroups WW and EE

Effect	D.F.	MS	F	Sig. of F
Group	1	303003	2.452	.126
Time	2	2136187	39.087	.0001
Group by Time	2	10570	.193	.825

6.4 Effects of Treatment on General Anxiety and Moderator Variables

As has been explained in section 3.6, general anxiety was not expected to be affected by treatment and the three general anxiety measures CSAQ.C, CSAQ.S, and A-Trait were administered with other purposes in mind. Nonetheless as explained at the beginning of this chapter it appears that general anxiety was affected by treatment.

Because these measures are similar and correlate highly or moderately with each other (c.f. section 3.6) and because two of them (CSAQ.C and CSAQ.S) were administered before treatment while the A-Trait was administered at the 4-week follow up, we can transform these measures into an equivalent index so that we can make pre-treatment and post treatment comparisons, using ANOVA and/or ANCOVA to test for between-group differences.

The main strategy involved in this transformation was to convert each subject's raw score on each scale into the percentage of the maximum possible score on that scale. Because the CSAQ.C correlates .67 ($p < .001$) while the CSAQ.S correlates only .40 ($p < .001$) with the A-Trait, CSAQ.C contribution towards the level of pre-treatment general anxiety was 45/100 (i.e., $.67^2$) while CSAQ.S contribution was 16/100 (i.e., $.40^2$).

I should remind the reader that both the CSAQ.C and the CSAQ.S total score ranges from 7 to 35 (c.f. section 4.4.1). 45/100 plus 16/100 is equal to 61/100. In order to make this

index comparable to the post-treatment one, which was simply a subject A-Trait score as a percentage of the total possible score on this measure, it was multiplied by 1.64 (note: 61/100 x 1.64 = 100/100).

So pre-treatment general anxiety levels were obtained as follows:

$$[(((\text{CSAQ.C} - 7)/28) * 100) * 0.45) + ((((\text{CSAQ.S} - 7)/28) * 100) * 0.16))] * 1.64$$

Note that because these two scales' minimum possible score was 7 while the maximum score was 35, a subject raw score would be decreased by 7 and then divided by 28 (35 minus 7) to obtain it as a fraction of the maximum possible score. The next step, multiplying this fraction by 100, transforms it into a percentage of the maximum possible score.

Obtaining post-treatment general anxiety levels was simpler,

$$((\text{A-Trait} - 20)/60) * 100$$

The reader is reminded that the possible range of scores on the A-Trait scale is 20 to 80 (c.f. section 4.4.2).

In this section I am going to report the effects of treatment on general anxiety levels measured by the index described above. A similar framework as for the previous section was adopted: firstly the results of the repeated measures ANOVA without the effects of moderator variables are

reported; secondly any significant moderator variables are identified and entered simultaneously in the repeated measures ANCOVA to establish the "salient" moderator variables; thirdly, only the salient moderator variables were entered in the ANCOVA which was finally reported.

6.4.1 Effects of treatment on general anxiety without accounting for moderator variables

Table 6.4.1.1 shows group general anxiety levels at pre-test and at follow-up while Figure 6.4.1.1 displays them graphically.

Insert Table 6.4.1.1 here

Insert Figure 6.4.1.1 here

The actual (post-treatment) A-Trait scores group means were 44.90, 41.24, 39.56, and 38.72 for groups 1, 2, 3, and 4 respectively. Such pattern of post-test group means between target and control groups was very similar to those obtained by Lent and Russell (1978).

TABLE 6.4.1.1

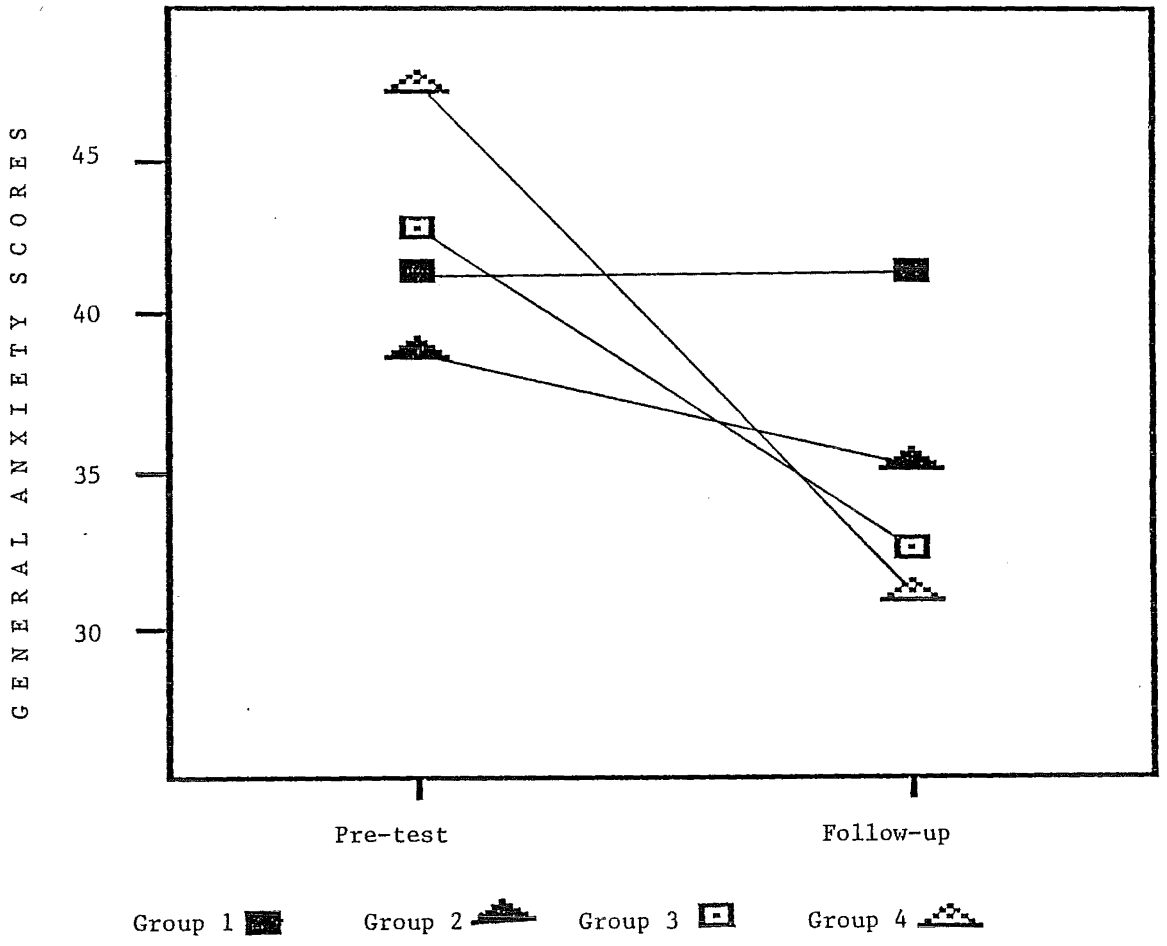
General anxiety levels at pre-test and follow-up
group means (standard deviations)

Group	n	pre-test	follow-up
1	9	41.40(13.94)	41.48(14.44)
2	8	38.88(12.77)	35.42(11.40)
3	11	42.84(15.16)	32.58(12.77)
4	7	47.74(16.00)	31.19(15.02)

FIGURE CAPTION

Figure 6.4.1.1: General Anxiety levels Group Means at Pre-test and Follow-up.

F I G U R E 6.4.1.1.


G E N E R A L A N X I E T Y S C O R E S
45
40
35
30
Pre-test
Follow-up
Group 1
Group 2
Group 3
Group 4

Insert Table 6.4.1.2 here

Table 6.4.12 shows the results of the repeated measures ANOVA. As we can see there was no significant effect for group but there were significant effects for time and group by time. A glance at the orthogonal contrasts for the latter effect reveals that group 4 decreased on general anxiety levels at a significantly higher rate than the rest ($t(1,35) = 2.341$, $p<.026$) while group 1 did so at a significantly lower rate ($t(1,35) = -2.113$, $p<.048$). Groups 2 and 3 did not produce significant orthogonal contrasts ($t(1,35) = -1.114$, $p<.274$ and $t(1,35) = .821$, $p<.418$ respectively).

6.4.2 Were individual groups improved on general anxiety?

Table 6.4.2.1 shows the results of analyses carried out on each individual group.

Insert Table 6.4.2.1 here

As we can see there was a highly significant effect for time when group 4 or group 3 is considered but such effect was far from significant when we consider group 2 and remotely so with respect to group 1.

TABLE 6.4.1.2

Repeated measures ANOVA on general anxiety

Effect	D.F.	MS	F	Sig. of F
Group	3	67	.211	.888
Time	1	934	12.739	.001
Group by Time	3	217	2.965	.047

TABLE 6.4.2.1

Repeated measures ANOVA on general anxiey for each of the four treatment groups taken individually.

Group	Effect	D.F.	MS	F	Sig. of F
1	Time	1	.03	.001	.978
2	Time	1	48	.397	.549
3	Time	1	580	7.972	.018
4	Time	1	958	15.244	.008

6.4.3 Looking for significant moderator variables

Twelve covariates were individually entered in the repeated measures ANCOVA and as many as six (E-NZ, AATD-F, ITA, TASEQ, SH, and Expectancy) produced significant regressions as can be seen in Table 6.4.3.1.

Insert Table 6.4.3.1 here

However when all six covariates were entered simultaneously, only two remained significant: ITA and SH (see Table 6.4.3.2); when only ITA and SH were entered simultaneously in the repeated measures ANCOVA ITA's Beta weight and significance level increased appreciably, unlike SH's.

Insert Table 6.4.3.2 here

Insert Table 6.4.3.3 here

TABLE 6.4.3.1

Covariates Beta weights and tests of significance

Covariate	Beta	T-Value	Sig. of T
E-NZ	-.395	-2.356	.025
AATD-F	.469	2.910	.007
ITA	.682	5.107	.001
TASEQ	-.480	-2.998	.005
SH	-.553	-3.632	.001
SE	-.097	- .531	.599
ET	-.239	-1.351	.187
ACHMOT	-.292	-1.672	.105
EWJ	.086	.457	.651
V-SCATT	-.218	-1.226	.230
N-SCATT	-.182	-1.015	.318
Expectancy	-.467	-2.893	.007

TABLE 6.4.3.2

Analysis of significant covariates when entered simultaneously in the repeated measures ANCOVA for general anxiety

Covariate	Beta	T-Value	Sig. of T
E-NZ	-.206	-1.410	.171
AATD-F	.122	.741	.466
ITA	.428	2.494	.020
TASEQ	.079	.430	.671
SH	-.357	-2.622	.015
Expectancy	-.095	- .680	.503

TABLE 6.4.3.3

Analysis of salient covariates in the repeated measures ANCOVA for general anxiety

Covariate	Beta	T-Value	Sig. of T
ITA	.555	4.269	.001
SH	-.353	-2.714	.011

6.4.4 Effects of treatment on general anxiety levels accounting for the two salient moderator variables.

Table 6.4.4.1 shows the results of the repeated measures ANCOVA with the two moderator variables as covariates.

Insert Table 6.4.4.1 here

Comparing this with Table 6.4.3.3 we see that there was no appreciable difference except for a highly significant regression. A look at the group by time orthogonal contrasts shows that they were essentially the same as when no covariates are entered in the analysis: ($t(1,35) = -2.054$, $p < .049$; $t(1,35) = -1.112$, $p < .275$; $t(1,35) = .798$, $p < .431$; $t(1,35) = 2.30$, $p < .027$, respectively for groups 1, 2, 3, and 4).

The covariates (ITA and SH) regression multiple R was .76 ($F = 19.493$, $p < .001$).

6.5 Effects of Treatment on Study and Exam Skills

In this section I am going to report the effects of treatment on study and exam skills measures. Firstly, the four groups were left intact and the data was analysed in the usual way; secondly the two groups that received study skills training (1 and 4) were combined into one group and the two

TABLE 6.4.4.1

Repeated measures ANCOVA on general anxiety with the two salient covariates: ITA and SH

Effect	D.F.	MS	F	Sig. of F
(Regression)	2	2816	19.493	.0001
Group	3	59	.411	.746
Time	1	934	12.739	.001
Group by Time	3	217	2.965	.047

groups that did not (2 and 3) were combined into another and the data again analysed as before. This step was taken because no significant difference was obtained when the groups were left intact.

6.5.1 Effects of treatment on study skills: groups intact

Table 6.5.1.1 shows group means and standard deviations at pre-test, post-test and follow-up for the three study skills measures while Tables 6.5.1.2 and 6.5.1.3 show the results of the MANOVA and the individual ANOVAs respectively.

Insert Table 6.5.1.1 here

Insert Table 6.5.1.2 here

Insert Table 6.5.1.3 here

We can see from the above tables that there was no group effect and a highly significant effect for time at both the multivariate and the univariate level. Moreover, two of the ANOVAs (for SH and ET) show a group by time effect that

TABLE 6.5.1.1

Study skills measures pre-test, post-test. and follow-up group means (standard deviation)

Group	n	Time of measurement	SH	SE	ET
1	9	pre-test	115.44(21.36)	32.56(6.82)	31.67(7.02)
		post-test	139.00(23.15)	36.33(6.58)	37.67(6.73)
		follow-up	150.78(22.06)	37.22(7.00)	37.67(5.05)
2	8	pre-test	133.13(28.06)	32.00(7.05)	34.38(7.63)
		post-test	134.88(22.22)	33.88(3.09)	36.88(7.77)
		follow-up	149.25(33.77)	34.63(4.81)	37.50(7.60)
3	11	pre-test	136.45(36.29)	34.18(5.74)	37.55(5.22)
		post-test	140.91(38.01)	35.27(5.85)	38.18(6.24)
		follow-up	148.27(43.64)	37.27(4.58)	38.55(4.59)
4	9	pre-test	135.33(23.15)	33.67(7.48)	32.89(5.60)
		post-test	165.22(24.99)	35.22(6.94)	39.11(6.41)
		follow-up	163.67(27.50)	38.56(7.83)	40.33(6.40)

TABLE 6.5.1.2

Repeated measures MANOVA on study skills measures

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	3	-1/2	14 1/2	9	.174	.679	.726
Time	2	0	31	6	.462	6.510	.001
Group by Time	3	1	31	18	.317	1.302	.190

TABLE 6.5.1.3

Repeated measures ANOVAs study skills measures

Measure	Effect	D.F.	MS	F	Sig. of F
SH	Group	3	1940	.874	.464
	Time	2	4825	18.046	.001
	Group by Time	6	551	2.062	.070
SE	Group	3	28	.327	.806
	Time	2	135	8.445	.001
	Group by Time	6	6	.388	.884
ET	Group	3	35	.375	.772
	Time	2	197	14.921	.001
	Group by Time	6	26	1.985	.080

approaches significance.

6.5.2 Did individual groups improve on study skills?

A MANOVA and three ANOVAs were carried out on each group taken individually to see which groups, if any, improved on these measures. Tables 6.5.2.1 and 6.5.2.2 show the results of MANOVA and the ANOVAs respectively.

Insert Table 6.5.2.1 here

Insert Table 6.5.2.2 here

We can see that group 2 and group 4 clearly improved significantly on these measures. All the analyses show highly significant effects when these two groups were considered, except for the SE scale where group 2 approached significance and group 4 was nearly significant. One of group 2 ANOVAs (for SH) was virtually significant and another (for ET) approached significance but the other two analyses were clearly not significant. Group 3 analyses were far from significant.

TABLE 6.5.2.1

Repeated measures MANOVA on the three study skills measures for each of the four groups taken individually

Group	Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
1	Time	2	0	6	6	.717	2.804	.028
2	Time	2	0	5	6	.603	1.871	.124
3	Time	2	0	8	6	.168	.580	.744
4	Time	2	0	6	6	.910	4.178	.004

TABLE 6.5.2.2

Repeated measures ANOVAs on the three study skills measures for each of the four treatment groups taken individually.

Measure	Group	Effect	D.F.	MS	F	Sig. of F
SH	1	Time	2	2913	13.289	.001
	2	Time	2	626	3.710	.051
	3	Time	2	392	1.146	.338
	4	Time	2	2548	8.256	.003
SE	1	Time	2	55	2.845	.088
	2	Time	2	15	1.527	.251
	3	Time	2	27	1.581	.230
	4	Time	2	56	3.398	.059
ET	1	Time	2	108	5.344	.017
	2	Time	2	22	2.976	.084
	3	Time	2	3	.186	.832
	4	Time	2	143	15.963	.001

6.5.3 Effects of treatment on study skills: groups combined

The layout of results in this section was identical to the previous one. Table 6.5.3.1 shows group means and standard deviations for the three measurements and Tables 6.5.3.2 and 6.5.3.3 show the results of the MANOVA and individual ANOVAs.

Insert Table 6.5.3.1 here

Insert Table 6.5.3.2 here

Insert Table 6.5.3.3 here

As in the previous section there was no significant effect for group but a highly significant effect for time at both the multivariate and the univariate levels. In addition there was a significant group by time effect on both the MANOVA and on two of the ANOVAs: for SH and for ET.

6.6 Effects of Treatment on Academic Performance

In this section I am going to report the results

TABLE 6.5.3.1

Study skills measures pre-test, post-test, and follow-up group means (standard deviations) for subjects receiving study skills training (groups 1 and 4) versus those who did not (Groups 2 and 3).

Group	n	Time of measurement	SH	SE	ET
1 + 4	18	pre-test	125.39(24.08)	33.11(6.97)	32.28(6.19)
		post-test	152.11(26.99)	37.78(6.58)	38.39(6.42)
		follow-up	157.22(25.08)	37.89(7.23)	39.00(5.76)
2 + 3	19	pre-test	135.06(32.26)	33.26(6.230	36.21(6.36)
		post-test	138.37(31.68)	34.68(4.82)	37.63(6.75)
		follow-up	148.68(38.75)	36.16(4.74)	38.11(5.87)

TABLE 6.5.3.2

Repeated measures MANOVA on study skills measures
(subjects receiving study skills vs those who did not)

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	1	1/2	15 1/2	3	.059	.689	.565
Time	2	0	33	6	.458	6.831	.001
Group by time	2	0	33	6	.216	2.791	.014

TABLE 6.5.3.3

Repeated measures ANOVAs on study skills measures (subjects receiving study skills training vs those who did not)

Measure	Effect	D.F.	MS	F	Sig. of F
SH	Group	1	490	.218	.643
	Time	2	4825	18.596	.001
	Group by Time	2	1396	5.381	.007
SE	Group	1	22	.265	.610
	Time	2	135	8.789	.001
	Group by Time	2	8	.554	.577
ET	Group	1	16	.176	.678
	Time	2	197	15.508	.001
	Group by Time	2	70	5.485	.006

pertaining to the effects of treatment on academic performance measures: FGA and TEA.

A MANOVA was not attempted because there was a substantial proportion of subjects whose data was not available on these measures. Moreover, data which are missing on one variable are available on the other. The consequence of this state of affairs was that very few subjects would be included in a MANOVA. Notwithstanding the above these two variables seemed to behave in different ways, as will be obvious from this and the next section.

For each of these two measures an ANOVA was carried out on all treated subjects without taking into account their scholastic ability scores. The ANOVA was in turn carried out only on those subjects whose percentile score was between 15 and 85 on the SCATT, V-SCATT, and N-SCATT.

As the FGA score but no TEA or scholastic ability scores were available for the control group (group 5) an ANOVA on all 5 groups on the FGA without taking into account ability was also be carried out.

6.6.1 FGA changes on treatment and control groups

Table 6.6.1.1 shows group means and standard deviations while Table 6.6.1.2 shows the results of the ANOVA.

Insert Table 6.6.1.1 here

TABLE 6.6.1.1

Pre-test and post-test FGA group means
(standard deviations)

Group	n	Pre-test	Post-test
1	5	64.20(11.54)	61.80(7.82)
2	5	59.60(8.05)	62.00(6.28)
3	7	54.71(8.20)	60.14(8.05)
4	6	55.00(9.40)	65.00(7.16)
5	11	57.91(8.75)	61.09(7.18)

Insert Table 6.6.1.2 here

There was a significant effect for time but not for group. The group by time effect was also not significant, but the group by time orthogonal contrasts revealed something of interest: group 4 increased at an almost significantly higher rate when compared with the other groups ($t(1,34) = -1.974$, $p < .058$) while group 1 increased at a lower rate, the orthogonal contrast approaching significance ($t(1,34) = 1.794$, $p < .083$).

The remaining three orthogonal contrasts were far from approaching significance: $t(1,34) = .387$, $p < .701$, $t(1,34) = -.568$, $p < .574$, $t(1,34) = .210$, $p < .835$ respectively for groups 2, 3, and 5.

6.6.2 FGA changes and scholastic ability ranges

Table 6.6.2.1 shows pre-test and post-test group means and standard deviations at the different scholastic ability ranges while Table 6.6.2.2 shows the results of each of the relevant ANOVAs.

Insert Table 6.6.2.1 here

TABLE 6.6.1.2

Repeated measures ANOVA on FGA: treatment groups and controls

Effect	D.F.	MS	F	Sig. of F
Group	4	48	.491	.742
Time	1	260	6.791	.014
Group by Time	4	57	1.485	.233

TABLE 6.6.2.1

Pre-test and post-test FGA group means (standard deviations) at different ranges of ability for the four treatment groups

Range	Group	n	Pre-test	Post-test
Percentile	1	2	70.00(1.41)	58.00(11.31)
SCATT	2	3	57.33(3.21)	63.67(3.51)
> 14	3	5	58.00(7.11)	61.00(9.70)
< 86	4	5	53.20(9.28)	63.20(6.20)

Range	Group	n	Pre-test	Post-test
V-SCATT	1	3	64.33(9.87)	59.33(8.33)
Percentile	2	4	56.25(3.40)	60.75(6.50)
> 14	3	5	58.00(7.11)	61.00(9.70)
< 86	4	5	53.20(9.28)	63.20(6.30)

Range	Group	n	Pre-test	Post-test
N-SCATT	1	1	69.00(0.00)	50.00(0.00)
Percentile	2	3	61.33(10.12)	63.67(3.51)
> 14	3	5	52.40(7.02)	60.60(9.71)
< 86	4	3	52.67(14.02)	66.33(6.81)

Insert Table 6.6.2.2 here

From the latter table we can see that there was clearly no significant eep1ect for group but that there was a significant effect for time in all but one of the ANOVAs (in the V-SCATT range). The group by time effect on the four ANOVAs reveals an interesting story: while there was no significant effect when ability range or V-SCATT range was considered the effect approached significance in the SCATT range and was within the significance level in the N-SCATT range.

Insert Table 6.6.2.3 here

Table 6.6.2.3 shows the orthogonal group by time contrasts corresponding to the ANOVAs reported in the previous table. Essentially these orthogonal contrasts paralleled the results of the ANOVAs. There were two groups that tended to be significantly different from the rest: group 4 in the positive direction and group 1 in the negative. Noteworthy also is group 3 which, although far from being significantly different in the first three analyses, tended to approach significance in the N-SCATT range ANOVA.

TABLE 6.6.2.2

Repeated measures ANOVA on FGA at different ranges of ability for the 4 treatment groups.

Range	Effect	D.F.	MS	F	Sig. of F
(All Subjects)	Group	3	63	.632	.603
	Time	1	209	5.008	.037
	Group by Time	3	74	1.784	.184

Range	Effect	D.F.	MS	F	Sig. of F
Scatt percentile > 14 < 86	Group	3	33	.449	.723
	Time	1	120	2.991	.112
	Group by Time	3	119	2.963	.079

Range	Effect	D.F.	MS	F	Sig. of F
V-SCATT percentile > 14 < 86	Group	3	19	.251	.859
	Time	1	136	2.906	.112
	Group by Time	3	72	1.528	.245

Range	Effect	D.F.	MS	F	Sig. of F
N-SCATT percentile > 14 < 86	Group	3	46	.360	.783
	Time	1	204	6.093	.039
	Group by Time	3	144	4.305	.044

TABLE 6.6.2.3

Group by Time orthogonal contrasts for repeated measures ANOVAs on FGA at different ranges of ability

Range	Group	n	T-Value	Sig. of T
(All subjects)	1	5	1.804	.087
	2	5	.420	.679
	3	7	- .506	.619
	4	6	-1.883	.075

Range	Group	n	T-Value	Sig. of T
SCATT	1	2	2.700	.021
percentile	2	3	- .018	.331
> 14	3	5	- .310	.763
< 86	4	5	-2.166	.053

Range	Group	n	T-Value	Sig. of T
V-SCATT	1	3	1.758	.102
percentile	2	4	- .329	.747
> 14	3	5	.032	.975
< 86	4	5	-1.768	.100

Range	Group	n	T-Value	Sig. of T
N-SCATT	1	1	3.158	.013
percentile	2	3	- .237	.819
> 14	3	5	-1.811	.108
< 86	4	3	-2.838	.022

6.6.3 TEA changes and scholastic ability ranges

Table 6.6.3.1 shows pre-test and post-test group means and standard deviations.

Insert Table 6.6.3.1 here

Obviously there was a decrease in TEA scores from pre-test to post-test. Table 6.6.3.2 shows the results of the ANOVAs at the four different ranges of ability.

Insert Table 6.6.3.2 here

It was apparent that there was no significant effect for group nor for group by time. Two of the analyses (ranges SCATT and N-SCATT) yielded a significant effect for time and a third (all subjects included) approached significance on the same effect.

6.6.5 Effects of general anxiety on academic performance Improvements following test anxiety treatment

As general anxiety levels were affected by treatment (c.f. section 6.4) the CSAQ and not the A-Trait was employed as a

TABLE 6.6.3.1

Pre-test and post-test TEA group means (standard deviations) at different ranges of ability

Range	Group	n	Pre-test	Post-test
(All subjects)	1	7	59.93(12.13)	55.46(16.40)
	2	6	55.28(9.28)	53.33(9.70)
	3	5	57.04(1.73)	53.94(7.26)
	4	4	65.42(1.83)	61.53(5.02)

Range	Group	n	Pre-test	Post-test
SCATT	1	3	57.13(16.42)	43.63(19.83)
percentile	2	5	55.16(10.36)	53.76(10.78)
> 14	3	5	57.04(1.73)	53.94(7.26)
< 86	4	2	63.90(0.85)	59.90(4.67)

Range	Group	n	Pre-test	Post-test
V-SCATT	1	4	54.60(14.33)	48.63(19.02)
percentile	2	5	55.04(10.35)	53.44(10.84)
> 14	3	5	57.04(1.73)	53.94(7.26)
< 86	4	2	63.90(0.85)	59.90(4.67)

Range	Group	n	Pre-test	Post-test
N-SCATT	1	1	63.40(0.00)	49.60(0.00)
percentile	2	3	60.77(4.52)	55.50(2.52)
> 14	3	3	57.77(1.97)	50.70(8.13)
< 86	4	3	64.97(1.94)	62.53(5.63)

TABLE 6.6.3.2

Repeated measures ANOVA on TEA at different ranges of ability

Range	Effect	D.F.	MS	F	Sig. of F
(All subjects)	Group	3	149	.882	.469
	Time	1	125	3.387	.082
	Group by Time	3	4	.100	.959

Range	Effect	D.F.	MS	F	Sig. of F
SCATT percentile > 14 < 86	Group	3	108	.531	.670
	Time	1	168	7.928	.017
	Group by Time	3	50	2.363	.127

Range	Effect	D.F.	MS	F	Sig. of F
V-SCATT percentile > 14 < 86	Group	3	97	.511	.682
	Time	1	96	2.000	.183
	Group by Time	3	7	.152	.926

Range	Effect	D.F.	MS	F	Sig. of F
N-SCATT percentile > 14	Group	3	94	3.029	.115
	Time	1	169	13.114	.011
	Group by Time	3	17	1.343	.346

criterion for low and high general anxiety in the analysis that follows. This was done to see whether low general anxiety subjects increased their academic performance more than their high counterparts as a result of these test anxiety treatments (c.f. section 3.6.2). Subjects who scored below this sample mean on the CSAQ (35) were assigned to the low general anxiety group and subjects who scored above the CSAQ mean were assigned to the high general anxiety group. Although the mean of a much larger sample would have been a more desirable criterion for the high and low general anxiety subdivision, such a mean was not available in the literature (c.f. Delmonte & Ryan, 1983; Schwartz, Davidson, & Goleman, 1978) and the mean for this sample had to be employed. Nonetheless, because a score of 35 was a little below half-way (42) between the possible range of scores on the CSAQ (c.f. section 4.4.1) it was likely to be similar to the mean for a large sample in view of the fact that the A-Trait scale, which was similar to the CSAQ, has a mean for a large sample of university students which was also a little below half-way between the possible range of scores (c.f. Spielberger, Gorsuch, & Lushene, 1970).

Tables 6.6.5.1 and 6.6.5.2 show FGA pre-test and post-test means (standard deviations) for the high and low general anxiety groups and resultant ANOVA.

Insert Table 6.6.5.1 here

TABLE 6.6.5.1

Low and high general anxiety FGA group means (standard deviations) at pre-test and post-test

Group	n	Pre-test	Post-test
Low general anxiety	8	60.75(9.41)	61.13(7.38)
High general anxiety	15	56.40(9.46)	62.73(7.23)

Insert Table 6.6.5.2 here

There was a significant effect for time but that the group by time effect was not significant. The group effect was far from significant ranges. Tables 6.6.5.3 and 6.6.5.4 show the FGA post-test and follow-up 12 group means (standard deviations) and relevant ANOVA. Group by time effects are far from significant ranges; the effect for group also was also not significant.

Insert Table 6.6.5.3 here

Insert Table 6.6.5.4 here

Tables 6.6.5.5 and 6.6.5.6 show TEA group means (standard deviations) and ANOVA as above for FGA. We can see that the time effect approached significance but that neither the group nor group by time did so.

Insert Table 6.6.5.5 here

TABLE 6.6.5.2

ANOVA on low and high general anxiety groups for FGA between pre-test and post-test

Effect	D.F.	MS	F	Sig. of F
Group	1	20	.200	.660
Time	1	209	4.752	.041
Group by Time	1	93	2.108	.161

TABLE 6.6.5.3

Low and general anxiety FGA group means (standard deviations) at post-test and follow-up

Group	n	post-test	follow-up 12
Low and general anxiety	12	59.17(7.11)	57.92(7.53)
High and general anxiety	16	62.25(8.56)	62.06(8.74)

TABLE 6.6.5.4

MANOVA on low and high general anxiety groups for FGA between post-test and follow-up 12

Effect	D.F.	MS	F	Sig. of F
Group	1	179	2.112	.158
Time	1	6	.124	.728
Group by time	1	4	.083	.776

TABLE 6.6.5.5

Low and high general anxiety TEA group means (standard deviations) at pre-test and post-test

Group	n	pre-test	post-test
Low and general anxiety	12	58.77(8.02)	54.58(11.68)
High and general anxiety	10	59.29(8.75)	56.91(10.67)

Insert Table 6.6.5.6 here

6.7 Effects of Treatment on Ability Measures

In this section I am going to report the results pertaining to the effects of treatment on the scholastic ability measures.

6.7.1 Data and results

Tables 6.7.1.1, 6.7.1.2, and 6.7.1.3 respectively show: pre-test and post-test group means and standard deviations, results of the repeated measures MANOVA, and results of the individual repeated measures ANOVAs.

Quite clearly all the effects were remote from significance levels.

Insert Table 6.7.1.1 here

Insert Table 6.7.1.2 here

TABLE 6.6.5.6

ANOVA on low and high general anxiety groups for TEA

Effect	D.F.	MS	F	Sig. of F
Group	1	22	.129	.724
Time	1	125	3.751	.067
Group by Time	1	9	.269	.610

TABLE 6.7.1.1

Pre-test and follow-up 12 V-SCATT and N-SCATT group means (standard deviation)

	V - S C A T T		
Group	n	Pre-test	Follow-up 12
1	6	33.67(6.22)	31.83(5.91)
2	5	29.00(4.69)	30.00(4.74)
3	9	32.89(7.88)	34.11(4.78)
4	6	34.50(6.66)	33.83(6.71)

	N - S C A T T		
Group	n	Pre-test	Follow-up 12
1	6	34.33(11.36)	32.50(8.38)
2	5	28.40(10.16)	28.20(10.26)
3	9	24.67(7.91)	26.67(8.66)
4	6	27.83(9.83)	29.00(10.20)

TABLE 6.7.1.2

MANOVA on scholastic ability measures

Effect	S -	M	N	D.F.	Value	Approx. F	Sig. of F
Group	2	0	9 1/2	6	.223	.922	.488
Time	1	0	9 1/2	2	.012	.131	.878
Group by Time	1	0	9 1/2	6	.132	.518	.791

Insert Table 6.7.1.3 here

6.8 Varieties of SCATT Ranges and Correlations between Test Anxiety Measures with Academic and Ability Tests Performance.

In this section I am going to report the correlations obtained between test anxiety measures and measures of academic performance and ability tests, as well as changes in test anxiety measures and academic and ability tests measures. The same correlations between the above variables were obtained and reported when (a) all subjects (with available data) were included, and when only subjects in the broad middle range (between the 15th and 85th percentile) of the (b) SCATT, (c) V-SCATT, and (d) N-SCATT distribution were included.

In line with test anxiety theory and previous research we would expect negative correlations between the AATD-F as well as the ITA with academic performance and ability tests performance. However, because of its nature, the TASEQ was expected to yield such correlations with a positive sign.

It was important to realize that because of the small number of available data many of these correlations - particularly those concerning the FGA and those in the SCATT, V-SCATT and N-SCATT ranges - would probably fail to reach significance. For the same reason the size of these correlations would tend to fluctuate considerably. The point

TABLE 6.7.1.3

Repeated measures ANOVAs on scholastic ability measures

Measure	Effect	D.F.	MS	F	Sig. of F
V-SCATT	Group	3	47	.795	.510
	Time	1	.02	.001	.974
	Group by Time	3	7	.395	.758
N-SCATT	Group	3	145	.869	.472
	Time	1	3	.270	.609
	Group by Time	3	10	.802	.506

I wish to make is that the tables printed in this secton show not the individual correlations that are of value but rather the trends in these correlations.

6.8.1 Correlations at pre-test, post-test, follow-up and follow-up 12

Table 6.8.1.1 shows the three test anxiety measures correlations with FGA, TEA, SCATT, V-SCATT, and N-SCATT at pre-test.

Insert Table 6.8.1.1 here

We can see that test anxiety measures tended to correlate with academic performance and with ability tests performance in the expected direction. Nonetheless, the most consistent trend apparent from the above table was that these correlations were higher and more significant in the N-SCATT range. Surprisingly, two of the AATD-F correlations with FGA were relatively large and significant but positive (instead of negative). Nonetheless, in line with the overall trend, such correlations turned in the expected direction (although without reaching significance) in the N-SCATT range.

Table 6.8.1.2 shows the follow-up 12 correlations between the same variables as in the previous table with the exception of TEA whose data was not collected.

TABLE 6.8.1.1

Correlations between pre-test anxiety measures and pre-test academic and ability measures at different ranges of aptitude (significance level

ALL SUBJECTS INCLUDED

	FGA (N=22)	TEA (N=22)	V-SCATT (N=22)	N-SCATT (N=37)	SCATT (N=37)
AATD-F	.03(.449)	-.12(.304)	-.33(.022)	-.46(.002)	-.49(.001)
ITA	-.15(.245)	-.46(.016)	-.34(.019)	-.08(.313)	-.22(.091)
TASEQ	.24(.137)	.25(.134)	.35(.017)	.32(.026)	.41(.006)

SUBJECTS WITH A SCATT percentile score >14 and <86

	FGA (N=15)	TEA (N=15)	V-SCATT (N=27)	N-SCATT (N=27)	SCATT (N=27)
AATD-F	.56(.015)	.15(.298)	-.15(.234)	-.40(.019)	-.42(.015)
ITA	-.03(.453)	-.48(.034)	-.23(.123)	.05(.395)	-.08(.343)
TASEQ	.13(.324)	.10(.362)	.14(.241)	.17(.193)	.24(.117)

SUBECTS WITH A V-SCATT percentile score >14 and <86

	FGA (N=17)	TEA (N=17)	V-SCATT (N=27)	N-SCATT (N=27)	SCATT (N=27)
AATD-F	.55(.011)	.17(.252)	.12(.272)	-.41(.017)	-.31(.055)
ITA	-.05(.426)	-.46(.032)	-.08(.355)	.06(.381)	.03(.439)
TASEQ	.10(.345)	.04(.435)	.06(.375)	.16(.207)	.18(.188)

SUBJECTS WITH A N-SCATT percentile score >14 and <86

	FGA (N=12)	TEA (N=10)	V-SCATT (N=20)	N-SCATT (N=20)	SCATT (N=20)
AATD-F	-.14(.338)	-.56(.047)	-.44(.025)	-.39(.046)	-.52(.009)
ITA	-.25(.218)	-.42(.111)	-.41(.038)	.04(.440)	-.28(.115)
TASEQ	.69(.006)	.45(.096)	.31(.093)	.00(.500)	.25(.146)

Insert Table 6.8.1.2 here

It was apparent that at this measurement there was a tendency for test anxiety measures not to correlate with performance measures or to correlate in the direction opposite to that expected. The highest and more significant correlations were in the N-SCATT range.

Table 6.8.1.3 shows the TEA post-test measurement correlations with test anxiety measures post-test and follow-up measurement. I should point out that post-test TEA was obtained from exams taken between the period after the test anxiety measures post-test and follow-up data collection. There are only two significant correlations; however as before we see that these correlations tended to be higher in the N-SCATT range.

Insert Table 6.8.1.3 here

TABLE 6.8.1.2

Correlations between test anxiety measures with academic and ability measures at follow-up 12 at different ranges of aptitude

	FGA (N=23)	V-SCATT (N=25)	N-SCATT (N=25)	SCATT (N=25)
ALL SUBJECTS INCLUDED				
AATD-F	.14(.255)	-.04(.424)	-.26(.107)	-.22(.144)
ITA	-.10(.331)	.09(.344)	.13(.278)	.14(.256)
TASEQ	-.01(.478)	.05(.412)	.25(.15)	.22(.148)
SUBJECTS WITH A SCATT percentile score >14 and <86				
	FGA (N=16)	V-SCATT (N=17)	N-SCATT (N=17)	SCATT (N=17)
AATD-F	-.13(.323)	.01(.491)	-.28(.136)	-.24(.180)
ITA	-.15(.294)	.04(.442)	.07(.304)	.08(.376)
TASEQ	-.01(.489)	.03(.453)	.14(.301)	.13(.304)
SUBJECTS WITH A V-SCATT percentile score >14 and <86				
	FGA (N=16)	V-SCATT (N=18)	N-SCATT (N=18)	SCCATT (N=18)
AATD-F	.03(.455)	.21(.203)	-.29(.121)	-.15(.277)
ITA	-.06(.414)	.14(.295)	.04(.444)	.15(.285)
TASEQ	-.18(.257)	-.17(.246)	.17(.250)	.06(.407)
SUBJECTS WITH A N-SCATT percentile score >14 and <86				
	FGA (N=14)	V-SCATT (N=14)	N-SCATT (N=14)	SCATT (N=14)
AATD-F	.23(.223)	.01(.485)	-.04(.443)	-.02(.471)
ITA	.38(.088)	.37(.096)	.48(.040)	.51(.030)
TASEQ	-.47(.047)	-.07(.408)	.22(.222)	.11(.36)

TABLE 6.8.1.3

Correlations between post-test and follow-up test anxiety measure with post-test TEA at different ranges of aptitude

	All subjects included	
	TEA with post-test test anxiety measures (n=37)	TEA with follow-up test anxiety measures (n=37)
AATD-F	-.07(.344)	-.21(.101)
ITA	-.25(.067)	-.29(.040)
TASEQ	-.08(.325)	.11(.251)

	Subjects with a SCATT percentile score >14 and <86	
	TEA with post-test test anxiety measures (n=27)	TEA with follow-up test anxiety measures (n=27)
AATD-F	.09(.321)	-.09(.330)
ITA	-.17(.195)	-.28(.080)
TASEQ	-.14(.246)	-.14(.249)

	Subjects with a V-SCATT percentile scores >14 and <86	
	TEA with post-test test anxity measures (n=27)	TEA with follow-up test anxiety measures (n=27)
AATD-F	.16(.209)	-.03(.435)
ITA	-.12(.274)	-.17(.204)
TASEQ	-.16(.214)	.03(.434)

	Subjects with a N-SCATTpercentile score >14 and <86	
	TEA with post-test test anxiety measure (n=20)	TEA with follow-up test anxiet measures (n=20)
AATD-F	-.19(.215)	-.24(.154)
ITA	-.24(.151)	-.44(.027)
TASEQ	.06(.395)	-.04(.427)

6.8.2 Test anxiety changes and performance changes.

In this section I am going to report correlations between changes in test anxiety measures and changes in academic performance and ability tests performance. A pre-test to post-test change was obtained by subtracting the pre-test score from the post-test score, and so on. Table 6.8.2.1 shows test anxiety changes (from pre-test to post-test and from pre-test to follow up) correlations with academic performance changes from pre-test to post-test.

Insert Table 6.8.2.1 here

A number of points are evident from this table. Firstly, correlations were higher and more significant between the pre-test to post-test (as opposed to the pre-test and follow-up) test anxiety measures change and the academic performance measures change. Secondly the correlations tended to be in the expected direction when considering FGA but in the direction opposite to the expected when considering TEA. In other words, greater decreases in test anxiety were associate with greater increases in FGA, but greater decreases in test anxiety was associated with greater decreases in TEA. Thirdly, and consistent with previous tables, correlations tended to be higher and more significant in the N-SCATT range.

Table 6.8.2.2 shows the correlations between test anxiety and performance changes from pre-test to follow-up 12.

TABLE 6.8.2.1

Correlations between changes (from pre-test to post-test and from pre-test to follow-up) with academic performance measures changes (from pre-test to follow-up) at different ranges of aptitude (significance level).

	Academic performance changes with test anxiety changes from pre-test to post-test		Academic performance changes with test anxiety changes from pre-test to follow-up	
	FGA(N=23)	TEA(N=22)	FGA(N=23)	TEA(N=22)
All subjects included				
AATD-F	-.12(.297)	.38(.040)	-.11(.310)	.16(.242)
ITA	-.17(.222)	.37(.046)	.03(.455)	.13(.282)
TASEQ	.05(.412)	-.50(.009)	.10(.333)	-.31(.081)
Subjects with a SCATT percentile score >14 and <86				
	FGA(N=15)	TEA(N=15)	FGA(N=15)	TEA(N=15)
AATD-F	-.19(.254)	.34(.108)	-.22(.213)	.07(.407)
ITA	-.24(.192)	.35(.098)	-.12(.331)	.00(.498)
TASEQ	-.03(.454)	-.51(.027)	.09(.375)	-.25(.189)
Subjects with a V-SCATT percentile score >14 and <86				
	FGA(N=17)	TEA(N=17)	FGA(N=17)	TEA(N=17)
AATD-F	-.15(.281)	.46(.031)	-.18(.242)	.28(.141)
ITA	-.22(.201)	.36(.079)	-.09(.362)	.20(.221)
TASEQ	-.02(.469)	-.52(.016)	.04(.443)	-.37(.071)
Subjects with a N-SCATT percentile score >14 and <86				
	FGA(N=12)	TEA(N=10)	FGA(N=12)	TEA(N=10)
AATD-F	-.55(.033)	.22(.274)	-.37(.116)	.06(.434)
ITA	-.51(.046)	.30(.200)	-.00(.494)	-.14(.347)
TASEQ	.34(.137)	-.44(.103)	.30(.137)	-.26(.238)

Insert Table 6.8.2.2 here

Similar observations can be made from this table. Firstly, when considering FGA the correlations were in the expected direction, but they were in the direction opposite to the expected when considering SCATT, V-SCATT, and N-SCATT. Secondly, both the size and significance of the correlations was higher in the N-SCATT range.

Table 6.8.2.3 shows the correlations between follow-up and follow-up 12 changes in test anxiety measures and FGA.

Insert Table 6.8.2.3 here

Once more FGA and test anxiety measures were in the expected direction and once more the size and significance of the correlations was higher in the N-SCATT range.

6.9 TASEQ Validation

In this section I am going to report results pertaining to the validity and reliability of the TASEQ.

TABLE 6.8.2.2

Correlations between test anxiety measures changes (from pre-test to follow-up 12) with academic and ability measures changes (from pre-test to follow-up 12) at different ranges of aptitude(significance level)

	FGA (N=14)	V-SCATT (N=25)	N-SCATT (N=25)	SCATT (N=25)
ALL SUBJECTS INCLUDED				
AATD-F	-.35(.113)	.19(.182)	.01(.472)	.15(.273)
ITA	-.24(.217)	-.06(.388)	.48(.009)	.21(.168)
TASEQ	.42(.065)	-.26(.105)	.08(.355)	-.15(.240)
	FGA (N=8)	**V-SCATT (N=17)**	**N-SCATT (N=17)**	**SCCATT (N=17)**
SUBJECTS WITH A SCATT percentile score >14 and <86				
AATD-F	-.41(.159)	.22(.195)	.16(.272)	.25(.168)
ITA	-.56(.074)	.17(.256)	.39(.060)	.31(.114)
TASEQ	.37(.186)	-.24(.180)	-.05(.426)	-.20(.218)
	FGA (N=10)	**V-SCATT (N=18)**	**N-SCATT (N=18)**	**SCATT (N=18)**
SUBECTS WITH A V-SCATT percentile score >14 and <86				
AATD-F	-.32(.187)	.13(.305)	-.01(.483)	.09(.363)
ITA	-.56(.059)	.16(.272)	.44(.038)	.32(.105)
TASEQ	.34(.167)	-.16(.260)	-.00(.497)	-.12(.322)
	FGA (N=8)	**V-SCATT (N=14)**	**N-SCATT (N=14)**	**SCATT (N=14)**
SUBJECTS WITH A N-SCATT percentile score >14 and <86				
AATD-F	-.57(.071)	.34(.119)	.05(.435)	.30(.147)
ITA	-.48(.112)	-.20(.249)	.56(.018)	.16(.290)
TASEQ	.60(.057)	-.43(.062)	-.01(.489)	-.36(.105)

TABLE 6.8.2.3

Correlations between test anxiety measures changes (from follow-up to follow-up 12) and academic performance changes (from follow-up to follow-up 12) at different ranges of aptitude (significance level)

ALL SUBJECTS

	FGA (n=23)
AATD-F	-.41(.027)
ITA	-.28(.094)
TASEQ	.35(.045)

SUBJECTS with a SCATT percentile score >14 and <86

	FGA (n=16)
AATD-F	-.38(.079)
ITA	-.15(.289)
TASEQ	.33(.106)

SUBJECTS with a V-SCATT percentile score > 14 and < 86

	FGA (n = 16)
AATD-F	-.51(.022)
ITA	-.21(.229)
TASEQ	.34(.099)

SUBJECTS with a N-SCATT percentile score >14 and <86

	FGA (n=14)
AATD-F	-.60(.015)
ITA	-.54(.002)
TASEQ	.29(.160)

6.9.1 Reliability of the TASEQ.

Cronbach's alpha coefficients were calculated on the questionnaires completed at pre-test, post-test, and follow up. The respective coefficients were .87, .97, and .96.

As we can see from Table 6.10.1.1 and Figure 6.10.1.1 in the next section TASEQ scores did not vary greatly between post-test and follow-up. We can therefore correlate TASEQ completed before the first exam and TASEQ completed before the last exam and treat it as a test-retest reliability of this questionnaire. The interval was approximately 10 days although it varied from subject to subject. The size of the correlation was .85($p < .001$), the number of subjects completing the questionnaire was 21 on both occasions.

6.9.2 Validity of TASEQ

The content validity of the TASEQ was assured in so far as its items were either derived from other test anxiety questionnaires or devised to relate to one's self-efficacy when confronted with situations known to produce anxiety in test anxious students. The items dealt with potential anxiety arousing from situations encountered during examinations but could easily be accommodated to deal with other forms of test-taking.

The correlations with academic and ability tests performance reported in Table 6.8.1.1 in the previous section establishes the TASEQ's concurrent validity. Evidence pertaining to the construct validity of the TASEQ was reported

on Tables 6.9.2.1, 6.9.2.2, 6.9.2.3, and 6.9.2.4.

Insert Table 6.9.2.1 here

Insert Table 6.9.2.2 here

Insert Table 6.9.2.3 here

Insert Table 6.9.2.4 here

Table 6.9.2.1 shows that the TASEQ correlated about as highly with general anxiety measures as the two test anxiety questionnaires employed in this study. Moreover the pattern of correlations between general anxiety and changing levels of test anxiety and self-efficacy was very similar to the equivalent pattern of correlations when the other two test anxiety measures were considered. Note that the CSAQ was administered at pre-test and the A-Trait at follow-up: the pattern of correlations with A-Trait was the converse of those with CSAQ measures.

Table 6.9.2.2 shows that TASEQ and its changes correlated

TABLE 6.9.2.1

Changing levels of test anxiety and correlations with general anxiety measures (significance levels)

	A A T D - F			
	Pre-test	Post-test	follow-up	follow-up 12
CSAQ.C	.31(.032)	.14(.202)	.26(.059)	.44(.012)
CSAQ.S	.05(.381)	.16(.178)	.08(.323)	.20(.160)
CSAQ	.21(.108)	.17(.155)	.20(.121)	.38(.027)
CSAQ-Trait	.27(.054)	.16(.169)	.24(.078)	.43(.015)
A-Trait	.46(.003)	.55(.001)	.67(.001)	.58(.001)

	I T A			
	Pre-test	Post-test	follow-up	follow-up 12
CSAQ.C	.47(.002)	.15(.188)	.12(.242)	.41(.020)
CSAQ.S	.31(.029)	.13(.214)	.09(.299)	.19(.179)
CSAQ	.48(.001)	.16(.169)	.12(.233)	.39(.026)
CSAQ-Trait	.45(.002)	.16(.166)	.12(.239)	.34(.043)
A-Trait	.49(.002)	.41(.007)	.52(.001)	.30(.070)

	T A S E Q			
	Pre-test	Post-test	follow-up	follow-up 12
CSAQ.C	-.46(.002)	-.13(.215)	-.21(.109)	-.48(.005)
CSAQ.S	-.29(.041)	-.10(.285)	.13(.219)	-.13(.252)
CSAQ	-.43(.004)	-.13(.216)	-.05(.394)	-.36(.031)
CSAQ-Trait	-.46(.002)	-.14(.208)	-.13(.216)	-.44(.011)
A-Trait	-.20(.128)	-.47(.002)	-.49(.001)	-.52(.003)

TABLE 6.9.2.2

Correlations between test anxiety measures (at post-test and follow-up) and test anxiety measures changes with item 1 of the PRQ and the full PRQ

	At post-test		At follow-up	
	SUBJ.EVAL	PRQ	SUBJ.EVAL	PRQ
AATD-F	-.08(.329)	-.05(.386)	-.26(.072)	-.13(.230)
ITA	.02(.461)	.09(.289)	-.13(.237)	.08(.319)
TASEQ	.34(.020)	.32(.028)	.28(.053)	.11(.274)

	With test anxiety changes from pre-test to post-test		With test anxiety changes from pre-test to follow-up	
	SUBJ.EVAL	PRQ	SUBJ.EVAL	PRQ
AATD-F	-.20(.119)	-.19(.125)	-.33(.030)	-.24(.086)
ITA	.07(.349)	.05(.395)	-.20(.127)	-.03(.428)
TASEQ	.39(.008)	.39(.009)	.24(.082)	.07(.338)

TABLE 6.9.2.3

Test anxiety measures intercorrelations (significance level) at the four measurements

	Pre-test (n = 37)		Post-test (n = 37)	
	ITA	TASEQ	ITA	TASEQ
AATD-F	.55(.001)	-.56(.001)	.72(.001)	-.48(.001)
ITA		-.53(.001)		-.29(.039)
TASEQ				

	Follow-up (n = 37)		Follow-up 12 (n = 37)	
	ITA	TASEQ	ITA	TASEQ
AATD-F	.61(.001)	-.58(.001)	.75(.001)	-.81(.001)
ITA		-.43(.004)		-.63(.001)
TASEQ				

TABLE 6.9.2.4

Test anxiety measures changes intercorrelations (significance level)

	Changes from pre-test to post-test (n=37)		Changes from pre-test to follow-up (n=37)	
	ITA	TASEQ	ITA	TASEQ
AATD-F	.60(.001)	-.66(.001)	.26(.061)	-.59(.001)
ITA		-.40(.007)		-.32(.025)
TASEQ				

	Changes from pre-test to follow-up 12 (n = 26)		Changes from post-test to follow-up (n = 37)	
	ITA	TASEQ	ITA	TASEQ
AATD-F	.51(.005)	-.82(.001)	.49(.001)	-.39(.009)
ITA		-.54(.002)		-.24(.080)
TASEQ				

	Changes from post to follow-up 12 (n = 2)		Changes from follow-up to follow-up 12 (n = 26)	
	ITA	TASEQ	ITA	TASEQ
AATD-F	.70(.001)	-.79(.001)	.69(.001)	-.57(.001)
ITA		-.53(.003)		-.62(.001)
TASEQ				

most highly and most significantly with the subjects own report of improvement (SUBJ.EVAL) as well as with the PRQ when compared with the ITA and the AATD-F.

Tables 6.9.2.3 and 6.9.2.4 respectively show intercorrelations between the TASEQ, the ITA and the AATD-F at the four different measurements and changes from one measurement to another.

Unlike the correlations between the ITA and the AATD-F, because of the nature of the TASEQ we would expect it to correlate negatively with the other two measures. We can see that the correlations were all in the expected directions and were virtually all highly significant.

Note how these correlations were first of a moderate magnitude, then decreased and finally increased to a moderately high level. Note also how the significance levels tended to decrease a little in the intermediate measurements but increased thereafter.

Finally the very high degree of internal consistency indicated by the previously reported Cronbach's Alpha was further evidence of the good construct validity of the TASEQ.

6.10 Self-Efficacy

In this section I am going to report the results pertaining to the self-efficacy aspect of this investigation.

6.10.1 Self-efficacy in managing one's anxiety (SEMA)

Table 6.10.1.1 shows TASEQ and P-TASEQ means and standard deviations when the available data from all students was pooled together.

Insert Table 6.10.1.1 here

A little over half of the subjects completed the TASEQ and P-TASEQ respectively before and after the first and last final exam. I should point out that the TASEQ at pre-test, post-test, follow-up and follow-up 12 referred to a hypothetical "scheduled exam or test" while it referred to the actual final exam on the other two occasions. There was a tendency to yield higher scores when referring to an actual exam. Figure 6.10.1.1 portrays the above means graphically. Note how closely TASEQ and P-TASEQ means and standard deviations matched each other.

Insert Figure 6.10.1.1 here

TASEQ and P-TASEQ correlated .77 ($p < .001$) at the first and .95 ($p < .001$) at the last exam.

The TASEQ correlated -.43 ($p < .028$) with the ITA at the first exam and -.67 ($p < .001$) at the last exam, the corresponding correlations between ITA and P-TASEQ were -.51

TABLE 6.10.1.1

TASEQ and P-TASEQ means (standard deviations) at the various measurement times

	At pre-test (n=37)	At post-test (n=37)
TASEQ	850(293)	1296(394)
P-TASEQ	-	-

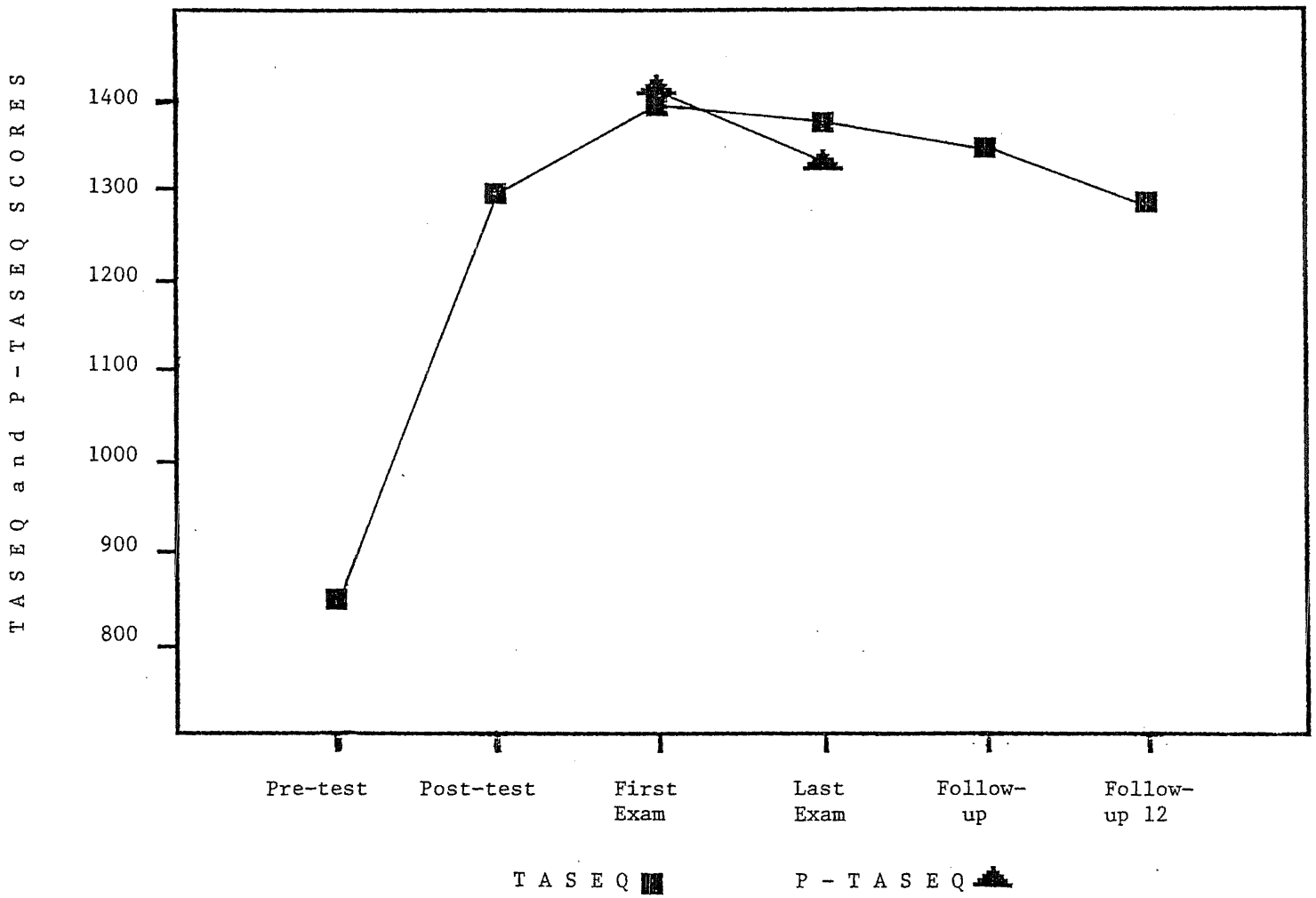
	Before/after subjects first exam (n=21)	Before/after subjects last exam (n=21)
TASEQ	1391(336)	1373(419)
P-TASEQ	1416(353)	1331(424)

	At follow-up (n=37)	At follow-up (n=27)
TASEQ	1347(366)	1288(430)
P-TASEQ	-	-

FIGURE CAPTION

Figure 6.10.1.1: TASEQ and PTASEQ changes over time.

F I G U R E 6.10.1.1.



(p< .010) and -.72 (p< .001). The P-TASEQ after the first exam correlated .83 (p< .001) with the P-TASEQ after the last exam while the corresponding correlation for the ITA was .62 (p< .001).

6.11 One Year Later

In this section I am going to report the results pertaining to the maintenance of treatment effects after twelve months following the completion of treatment.

6.11.1 Follow-up 12 on test anxiety measures: no moderator variables taken into account

5.11.1.1 shows the means and standard deviations for the three test anxiety measures at follow-up and follow-up 12.

Insert Table 6.11.1.1 here

Note that although 27 subjects were successfully contacted at the follow-up 12, because of incompleted questionnaires only 25 could be included in the analysis. Table 6.11.1.2 shows the results of the relevant MANOVA. At the multivariate level none of the effects even approached significance.

TABLE 6.11.1.1

Test anxiety measures follow-up and follow-up 12
group means (standard deviations)

Group	n	Time of measurement	AATD-F	ITA	TASEQ
1	6	Follow-up	3.33(15.02)	38.17(72.70)	1238.33(308.18)
		Follow-up 12	.33(16.07)	34.17(13.50)	1301.67(392.70)
2	5	Follow-up	5.20(15.72)	34.20(10.33)	1396.00(324.74)
		Follow-up 12	3.20(12.54)	34.00(12.75)	1278.00(478.40)
3	8	Follow-up	7.00(7.76)	38.38(11.19)	1323.75(433.46)
		Follow-up 12	4.88(11.04)	37.88(10.20)	1391.25(451.11)
4	6	Follow-up	-1.33(16.24)	29.00(6.90)	1501.67(444.27)
		Follow-up 12	4.33(12.31)	33.67(11.00)	1161.67(534.77)

Insert Table 6.11.1.2 here

The univariate analyses results were presented on Table 6.11.1.3.

Insert Table 6.11.1.3 here

We can see that there was only one significant effect, the group by time effect on the TASEQ. An examination of the orthogonal contrasts reveals that group 4 was the only one to differ significantly ($t(1,21) = 2.945$, $p < .008$). Group 3 contrast approached significance ($t(1,21) = -1.866$, $p < .076$) but neither group 1 nor group 2 did so ($t(1,21) = -1.655$, $P < .113$ and $t(1,21) = .388$, $p < .702$ respectively).

6.11.2 Changes of test anxiety measures and moderator variables

In this section the same analyses as were outlined in the previous section were reported except that the four moderator variables identified in section 6.3.2 were taken into account. Tables 6.11.2.1 and 6.11.2.2 show respectively the results of the MANCOVA and of the ANCOVAs.

TABLE 6.11.1.2

Repeated measures MANOVA on test anxiety measures between follow-up and follow-up 12

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	3	-1/2	8 1/2	9	.200	.501	.868
Time	1	1/2	8 1/2	3	.129	.936	.443
Group by Time	3	-1/2	8 1/2	9	.432	1.76	.326

TABLE 6.11.1.3

Repeated measures ANOVAs on individual test anxiety measures between follow-up and follow-up 12

Measure	Effect	D.F.	MS	F	Sig. of F
AATD-F	Group	3	60	.197	.897
	Time	1	2	.054	.818
	Group by Time	3	50	1.108	.368
ITA	Group	3	113	.518	.674
	Time	1	.02	.001	.980
	Group by Time	3	38	1.224	.326
TASEQ	Group	3	18367	.054	.983
	Time	1	58482	1.894	.183
	Group by Time	3	117975	3.816	.025

Insert Table 6.11.2.1 here

Insert Table 6.11.2.2 here

We can see that, except from obtaining significant regressions, both at a multivariate at a univariate levels, there was little difference when comparing with the results reported in the previous section. The orthogonal contrasts on the TASEQ group by time effect were also similar - group 1: $t(1,21) = -1.481$, $p < .154$; group 2: $t(1,21) = .338$, $p < .739$; group 3: $t(1,21) = -1.841$, $p < .080$; group 4: $t(1,21) = 2.766$, $p < .012$.

Table 6.11.2.3 shows the analyses of the univariate multiple regressions.

Insert Table 6.11.2.3 here

If we compare this with Table 6.3.2.10 we can see that the size of the multiple R increased appreciably for the ITA and the TASEQ while no appreciable change occurred on the AATD-F. Moreover, the significance level of the three regressions decreased (c.f. Table 6.11.2.2). Another interesting change was the contribution of the individual covariates to the size

TABLE 6.11.2.1

Repeated measures MANCOVA on test anxiety measures between follow-up and follow-up 12 with the effects of the salient moderator personality variables
(E-NZ, V-SCATT, CSAQ-Trait, ACHMOT) controlled for.

Effect	S	M	N	D.F.	Value	Approx.F	Sig. of F
(Regression)	3	0	6 1/2	12	1.007	2.146	.030
Group	3	1/2	6 1/2	9	.317	.669	.733
Time	1	1/2	8 1/2	3	.129	.936	.443
Group by Time	3	-1/2	8 1/2	9	.432	1.176	.326

TABLE 6.11.2.2

Repeated measures ANCOVAs on test anxiety measures between follow-up and follow-up 12 with the effects of the salient moderator personality variables (E-NZ, V-SCATT, CSAQ-Trait, ACHMOT) controlled for.

Measure	Effect	D.F.	MS	F	Sig. of F
AATD-F	(Regression)	4	795	4.142	.016
	Group	3	85	.445	.724
	Time	1	2	.054	.818
	Group by Time	3	50	1.100	.368
ITA	(Regression)	4	529	3.632	.026
	Group	3	159	1.092	.379
	Time	1	.02	.001	.980
	Group by Time	3	38	1.224	.326
TASEQ	(Regression)	4	899925	4.383	.013
	Group	3	186804	.910	.457
	Time	1	58482	1.894	.183
	Group by Time	3	117795	3.816	.025

TABLE 6.11.2.3

Analyses of the four salient moderator variables univariate multiple regression on test anxiety changes between follow-up and follow-up 12

Dep. Variable (Multiple R)	Covariate	Beta	T-Value	Sig. of T
AATD-F	E-NZ	- .352	-1.918	.072
(.703)	CSAQ-Trait	.257	1.368	.189
	V-SCATT	-.225	-1.275	.220
	ACHMOT	- .338	-1.932	.070
ITA	E-NZ	.024	.128	.899
(.679)	CSAQ-Trait	.455	2.352	.031
	V-SCATT	- .122	- .669	.513
	ACHMOT	- .425	-2.358	.031
TASEQ	E-NZ	.101	.559	.538
(.713)	CSAQ-Trait	- .371	-2.001	.061
	V-SCATT	.402	2.309	.034
	ACHMOT	.282	1.636	.120

of the regression. Again comparing Tables 6.11.2.3 and 6.3.2.10 we note that E-NZ and V-SCATT Beta weights and their significance decreased on all those test anxiety measures while the reverse happened with respect to the CSAQ-Trait and ACHMOT. The only exception to the above trend was the decrease in the significance level for ACHMOT on the AATD-F regression.

6.11.3 Follow-up 12 on SE and ET

As before (c.f. section 6.5) the study skills data were first analysed, leaving groups intact. The the two groups that received study skills training were combined into one, while the two groups that did not receive any study skills training (2 and 3) were combined into another. The data were then again analysed.

Insert Table 6.11.3.1 here

Insert Table 6.11.3.2 here

TABLE 6.11.3.1

SE and ET follow-up and follow-up 12 intact groups means (standard deviations)

Group	n	Time of measurement	SE	ET
1	7	Follow-up	37.86(6.01)	37.43(5.19)
		Follow-up 12	34.43(4.65)	36.00(2.58)
2	5	Follow-up	35.00(6.04)	37.20(9.09)
		Follow-up 12	35.40(4.16)	38.40(7.30)
3	8	Follow-up	38.63(4.27)	40.25(3.45)
		Follow-up 12	37.38(4.93)	41.13(3.64)
4	7	Follow-up	38.00(8.37)	40.14(7.17)
		Follow-up 12	37.00(6.98)	40.14(7.24)

TABLE 6.11.3.2

Repeated measures MANOVA on SE and ET
between follow-up and follow-up 12: groups intact

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	2	0	10	6	.123	.501	.804
Time	1	0	10	2	.089	1.069	.361
Group by Time	2	0	10	6	.136	.559	.760

Insert Table 6.11.3.3 here

Tables 6.11.3.1, 6.11.3.2, and 6.11.3.3 respectively show group means, MANOVA and ANOVAs for the two study skills scales measured at follow-up 12: SE and ET. As we can see none of the effects even approached significance, either at the multivariate or the univariate level.

The results of the analysis when the groups were combined as previously were reported on Tables 6.11.3.4, 6.11.3.5, and 6.11.3.6. As above, when the groups remained intact none of the effects even approached significance.

Insert Table 6.11.3.4 here

Insert Table 6.11.3.5 here

Insert Table 6.11.3.6 here

6.11.4 Follow-up 12 on academic performance: FGA

The format of this section will parallel that of section

TABLE 6.11.3.3

Repeated measures ANOVAs on SE and ET between follow-up and follow-up 12: groups intact

Measure	Effect	D.F.	MS	F	Sig. of F
SE	Group	3	20	.368	.777
	Time	1	28	2.209	.151
	Group by Time	3	8	.605	.619
ET	Group	3	50	.848	.482
	Time	1	.2	.022	.884
	Group by Time	3	5	.592	.627

TABLE 6.11.3.4

SE and ET follow-up and follow-up 12 combined groups means (standard deviations)

Group	n	Measurement	SE	ET
1 + 4	14	Follow-up	37.93(7.00)	38.79(6.18)
		Follow-up 12	35.71(5.85)	38.07(5.65)
2 + 3	13	Follow-up	37.23(5.12)	39.08(6.08)
		Follow-up 12	36.62(4.57)	40.08(5.24)

TABLE 6.11.3.5

Repeated measures MANOVA on SE and ET
between follow-up and follow-up 12: groups combined

Effects	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	1	0	11	2	.020	.242	.787
Time	1	0	11	2	.085	1.119	.343
Group by Time	1	0	11	2	.074	.962	.396

TABLE 6.11.3.6

Repeated measures ANOVAs on SE and ET between follow-up and follow-up 12: groups combined

Measure	Effect	D.F.	MS	F	Sig. of F
SE	Group	3	.1	.003	.960
	Time	1	28	2.288	.143
	Group by Time	3	9	.700	.411
ET	Group	3	18	.296	.591
	Time	1	.2	.023	.880
	Group by Time	3	10	1.372	.252

6.6.1 where the effects of treatment on FGA were reported. Note that FGA levels were available for all subjects at post-test (1983), but many did not enrol the following year (1984) and even fewer were enrolled the previous year (1982); moreover, there was very little overlap between subjects enrolled in 1982 and 1984. As a result of the above, data analysed in this section come from mostly different subjects when compared to the analyses of section 6.6.1.

Table 6.11.4.1 shows the 5 group means and standard deviations while Table 6.11.4.2 shows the results of the ANOVA.

Insert Table 6.11.4.1 here

Insert Table 6.11.4.2 here

There was no evidence of any significant effect from the analysis just reported.

6.11.5 Follow-up 12 on FGA and scholastic ability ranges

Table 6.11.5.1 shows follow-up and follow up 12 means and standard deviations for the four treatment groups at the

TABLE 6.11.4.1

Follow-up and follow-up 12
FGA group means (standard deviations)

Group	n	Follow-up	Follow-up 12
1	5	60.20(7.22)	58.80(6.38)
2	7	56.71(7.65)	54.71(8.06)
3	8	61.25(9.47)	64.25(4.20)
4	8	64.75(6.41)	62.13(10.88)
5	17	59.94(7.71)	59.47(10.22)

TABLE 6.11.4.2

Repeated measures ANOVA on FGA between follow-up and follow-up 12: treatment groups plus controls

Effect	D.F.	MS	F	Sig. of F
Group	4	142	1.363	.264
Time	1	8	.208	.651
Group by Time	4	19	.531	.713

different scholastic ability ranges while Table 6.11.5.1 shows the results of each of the relevant ANOVAs.

Insert Table 6.11.5.1 here

Insert Table 6.11.5.2 here

The results we obtained in this section parallel those of section 6.6.2: group 4 was discovered to be superior in the N-SCATT broad middle range, the difference being almost significant. The effect for group was almost significant on the N-SCATT range even though a large (12 out of 28) proportion of subjects were lost through the failure of their N-SCATT score to fall above the 14th and below the 86th percentiles. Such proportion of lost subjects was larger than the other two ability ranges.

Insert Table 6.11.5.3 here

By examining Table 6.11.5.3 we see that the group effect approaching significance when all subjects were included was due to group 2 being significantly below the others. In the N-SCATT range, however, the group effect being almost significant was due to group 4 which produced a significant

TABLE 6.11.5.1

Follow-up and follow-up 12 FGA performance (standard deviations) at diiferent ranges of ability for the four treatment groups

Range	Group	n	Follow-up	Follow-up 12
SCATT	1	2	57.00(12.73)	54.50(6.36)
percentile	2	6	57.50(8.07)	53.67(8.29)
> 14	3	6	62.33(10.95)	65.00(4.29)
< 86	4	6	63.83(6.05)	59.50(11.36)

Range	Group	n	Follow-up	Follow-up 12
V-SCATT	1	3	58.67(9.45)	57.00(6.25)
percentile	2	6	57.50(8.07)	53.67(8.29)
> 14	3	5	60.00(10.44)	63.80(3.49)
< 86	4	5	62.80(6.14)	59.20(12.68)

Range	Group	n	Follow-up	Follow-up 12
N-SCATT	1	0	-	-
percentile	2	4	58.25(5.06)	52.75(9.54)
> 14	3	7	61.29(10.23)	64.57(4.43)
< 86	4	5	66.40(5.13)	64.20(8.70)

TABLE 6.11.5.2

Repeated measures ANOVAs on FGA between follow-up and follow-up 12 at different ranges of ability: the four treatment groups only

Range	Effect	D.F.	MS	F	Sig. of F
(All subjects)	Group	3	183	2.392	.094
	Time	1	6	.121	.731
	Group by time	3	26	.536	.662

Range	Effect	D.F.	MS	F	Sig. of F
SCATT	Group	3	167	1.923	.167
percentile	Time	1	36	.575	.459
> 14	Group by time	3	31	.488	.695
<86					

Range	Effect	D.F.	MS	F	Sig. of F
V-SCATT	Group	3	91	1.111	.375
percentile	Time	1	19	.294	.596
> 14	Group by time	3	33	.498	.689
< 86					

Range	Effect	D.F.	MS	F	Sig. of F
N-SCATT	Group	2	229	3.526	.060
percentile	Time	1	3	.062	.808
> 14	Group by time	2	54	1.059	.375
< 86					

TABLE 6.11.5.3

Orthogonal contrasts for the group effect from repeated measures ANOVAs on FGA between follow-up and follow-up 12 at different ranges of ability

Range	Group	n	T-Value	Sig. of T
(All subjects)	1	5	- .371	.714
	2	7	-2.276	.032
	3	8	1.230	.231
	4	8	1.582	.127

Range	Group	n	T-Value	Sig. of T
SCATT	1	2	.928	.367
percentile	2	6	-1.425	.173
> 14	3	6	1.790	.092
< 86	4	6	.994	.335

Range	Group	n	T-Value	Sig. of T
V-SCATT	1	3	- .413	.686
percentile	2	6	-1.464	.164
> 14	3	5	1.117	.282
< 86	4	5	.760	.459

Range	Group	n	T-Value	Sig. of T
N-SCATT	1	0	-	-
percentile	2	4	- .878	.396
> 14	3	7	2.080	.058
< 86	4	5	2.564	.024

positive orthogonal contrast and group 3 which produced an almost significant one. Note how Group 1 was lost from the analysis as were all subjects. N-SCATT scores failed to be within the desired range, thereby decreasing the relevant degrees of freedom by one third.

6.12 Long Term Treatment Effects on Test Anxiety.

In this section test anxiety levels before treatment and at the follow-up 12 were examined to see if long term effects of treatment on test anxiety differed between treatments and to discover the subject variables that significantly moderated long term treatment effects.

6.12.1 Long term treatment effects: no moderator variables taken into account

Table 6.12.1.1 shows the appropriate group means and standard deviations while Tables 6.12.1.2 and 6.12.1.3 respectively show the results of the MANOVA and ANOVAs.

Insert Table 6.12.1.1 here

Insert Table 6.12.1.2 here

TABLE 6.12.1.1

Test anxiety measures pre-test and follow-up 12
group means (standard deviations)

Group	n	time of measurement	AATD-F	ITA	TASEQ
1	6	Pre-test	16.17(7.14)	44.33(12.53)	946.67(283.60)
		Follow-up 12	.33(16.07)	34.17(13.50)	1301.67(392.70)
2	5	Pre-test	19.40(4.62)	48.00(10.07)	696.00(321.37)
		Follow-up 12	3.20(12.54)	34.00(12.75)	1278.00(478.40)
3	8	Pre-test	19.25(8.80)	47.00(14.45)	827.50(326.79)
		Follow-up 12	4.88(11.04)	37.88(10.20)	1391.25(451.11)
4	6	Pre-test	15.83(7.11)	43.67(7.28)	721.67(257.01)
		Follow-up 12	4.33(12.31)	33.67(11.00)	1161.67(534.77)

TABLE 6.12.1.2

Repeated measures MANOVA on test anxiety measures between pre-test and follow-up 12

Effect	S	M	N	D.F.	Value	Approx. F	Sig. of F
Group	3	-1/2	8 1/2	9	.245	.624	.772
Time	1	1/2	8 1/2	3	.697	14.535	.001
Group by Time	3	-1/2	8 1/2	9	.249	.633	.764

Insert Table 6.12.1.3 here

We can see that at both the multivariate and univariate levels there was a highly significant effect for time. However, the group and group by time effects failed to even approach significance. Figures 6.12.1.1, 6.12.1.2 and 6.12.1.3 show each of the test anxiety measures group means graphically.

Insert Figure 6.12.1.1 here

Insert Figure 6.12.1.2 here

Insert Figure 6.12.1.3 here

6.12.2 Looking for significant long term moderator variables

Three steps were followed in this section which were very similar to those followed on section 6.3.2. Firstly, E-NZ, CSAQ-Trait, V-SCATT, and ACHMOT plus N-SCATT and SH were

TABLE 6.12.1.3

Repeated measures ANOVAs on test anxiety measures between pre-test and follow-up 12

Measure	Effect	D.F.	MS	F	Sig. of F
AATD-F	Group	3	36	.216	.884
	Time	1	2592	46.904	.001
	Group by time	3	13	.238	.869
ITA	Group	3	41	.201	.895
	Time	1	1394	18.647	.001
	Group by time	3	13	.175	.912
TASEQ	Group	3	100925	.478	.701
	Time	1	2971922	31.260	.001
	Group by time	3	35006	.368	.777

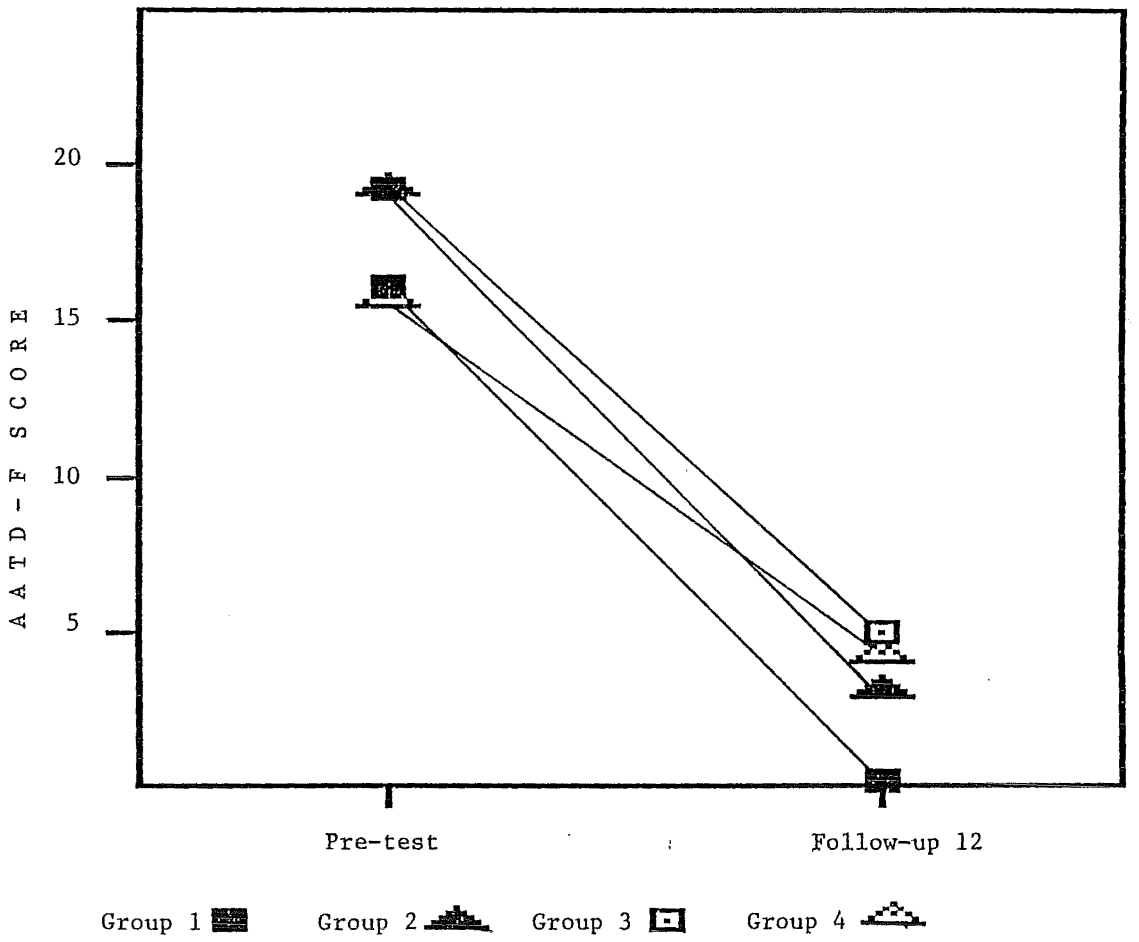
FIGURE CAPTIONS

Figure 6.12.1.1: AATD-F Group Means at Pre-test and Follow-up 12.

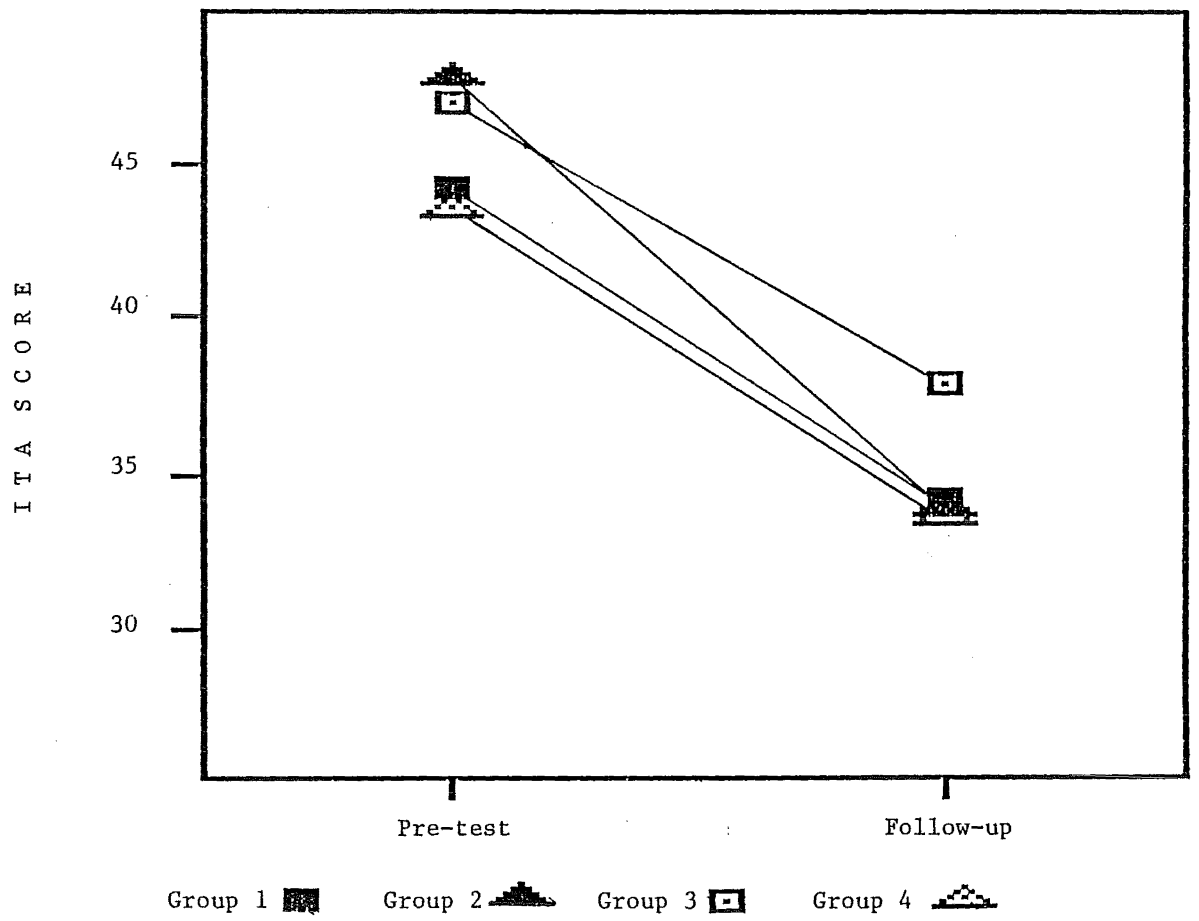
Figure 6.12.1.2: ITA Group Means at Pre-test and Follow-up 12.

Figure 6.12.1.3: TASEQ Group Means at Pre-test and Follow-up 12.

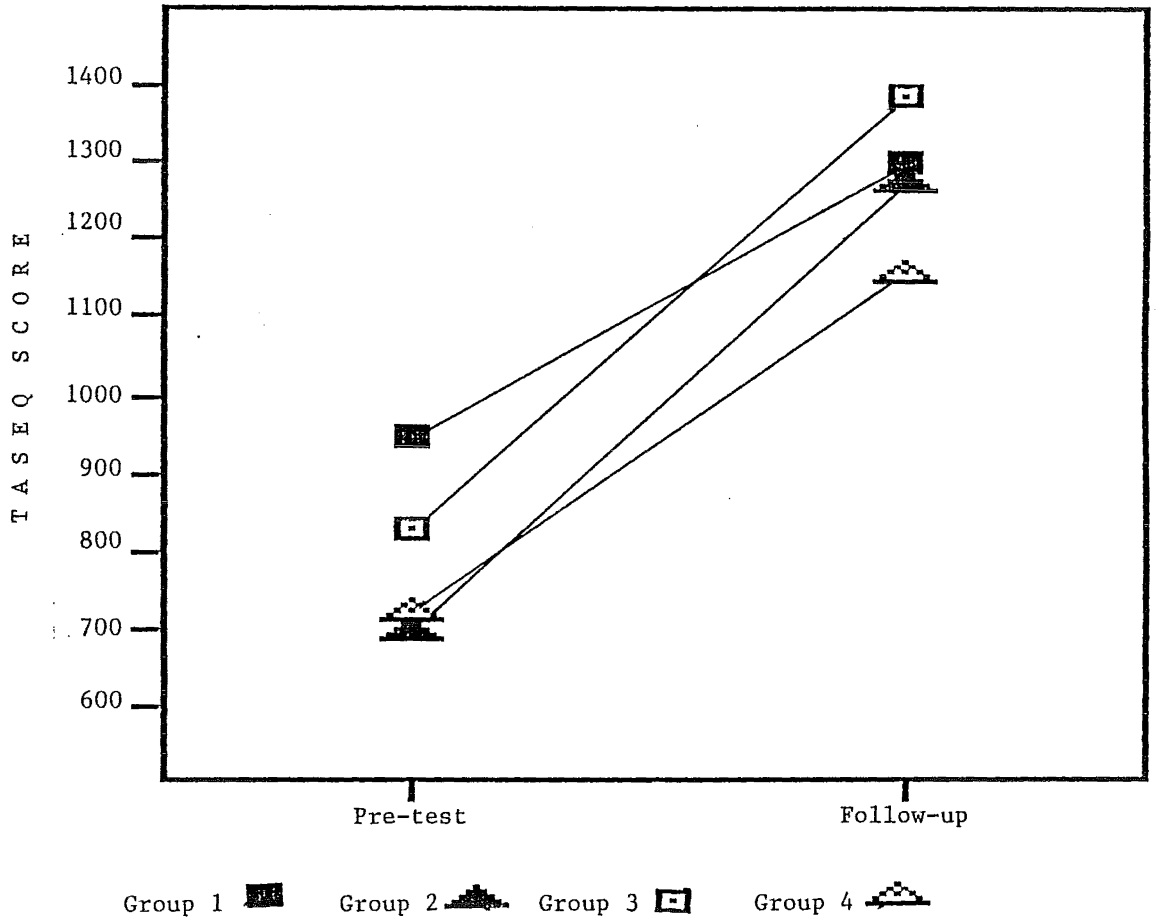
F I G U R E 6.12.1.1.



F I G U R E 6.12.1.2.



F I G U R E 6.12.1.3.



entered in the MANCOVA; the first four because they had already been found to be significant covariates in the section noted above and the other two because they individually produced a near significant regression when entered in the MANCOVA on test anxiety measures (pre-test, post-test, and follow-up) levels (c.f. Table 6.3.2.1). Secondly, any of the above covariates which did not appear to contribute significantly to any of the univariate regressions (i.e., for AATD-F, ITA, and TASEQ) were excluded and the MANCOVA was again carried out to see if the regression was substantially affected.

Once the selected covariates among the above six were identified the remaining five (EWJ, SE, ET, Expectancy, and CSAQ.S) were in turn entered in the same fashion as in section 6.3.2 to see if any of them substantially enhanced the size and significance of the regressions (i.e., personality variables regressions on each of the test anxiety measures ANCOVAs).

Table 6.12.2 shows the analyses of the six selected covariates multiple regressions.

Insert Table 6.12.2.1 here

Because the degrees of freedom were limited, not all covariates that fail to reach significance were excluded. Those that clearly appeared not to contribute either substantially or significantly to any of the three regressions were excluded. From the above table it appears that E-NZ and

TABLE 6.12.2.1

Analyses of the six significant covariates (E-NZ, CSAQ-Trait, V-SCATT, ACHMOT, and N-SCATT) multiple regressions on long term treatment effects on test anxiety

Dep. Variable (Mul. R, F, Sig. of F)	Covariate	Beta	T-Value	Sig. of T
AATD-F	E-NZ	-.144	-.722	.481
(R=.763,	CSAQ-Trait	.170	.786	.444
F(6,15)=3.483,	V-SCATT	-.211	-1.120	.280
p<.023)	ACHMOT	-.296	-1.673	.115
	SH	-.184	-.822	.424
	N-SCATT	-.335	-1.668	.116
ITA	E-NZ	-.016	-.095	.925
(R=.832,	CSAQ-Trait	.742	3.998	.001
F(6,15)=5.604,	V-SCATT	-.245	-1.508	.152
p<.003)	ACHMOT	-.197	-1.300	.215
	SH	.043	.221	.828
	N-SCATT	.099	.575	.574
TASEQ	E-NZ	.244	1.517	.150
(R=.854,	CSAQ-Trait	-.304	-1.745	.101
F(6,15)=6.715,	V-SCATT	.390	2.563	.022
p<.001)	ACHMOT	.085	.593	.562
	SH	.232	1.284	.219
	N-SCATT	.142	.880	.393

SH did not contribute either substantially or significantly to any of the three regressions.

Another MANCOVA was carried out which was identical to the previous one except that neither E-NZ nor SH were included as covariates. Table 6.12.2.2 shows the analyses of the three regressions when only the four selected moderator variables were included in the MANCOVA.

Insert Table 6.12.2.2 here

We can see that now all four yielded significant T-values on one or more of the regressions. Moreover, the size of the multiple Rs did not substantially decrease but their significance was further enhanced. This of course indicates that E-NZ and SH were not significant covariates.

The next step involved entering each of the remaining five covariates one by one into the same MANCOVA with the four selected moderator variables (ACHMOT, CSAQ-Trait, V-SCATT, N-SCATT) as covariates.

Insert Table 6.12.2.3 here

Insert Table 6.12.2.4 here

TABLE 6.12.2.2

Analyses of the four selected covariates
(N-SCATT, CSAQ-Trait, V-SCATT, and ACHMOT) multiple regressions
on long term treatment effects on test anxiety.

Dep. Variable (Mul. R, F, Sig. of T)	Covariate	Beta	T-Value	Sig. of T
AATD-F	N-SCATT	-.359	-2.108	.050
(R=.742,	CSAQ-Trait	.304	1.808	.088
F(4,17)=5.27,	V-SCATT	-.243	-1.410	.177
p<.006)	ACHMOT	-.347	-2.112	.050
ITA	N-SCATT	.083	.585	.566
(R=.831,	CSAQ-Trait	.725	5.192	.001
F(4,17)=9.472,	V-SCATT	-.230	-1.613	.125
p< .001)	ACHMOT	-.188	-1.380	.186
TASEQ	N-SCATT	.200	1.355	.193
(R=.813,	CSAQ-Trait	-.490	-3.359	.004
F(4,17)=8.291,	V-SCATT	.421	2.814	.012
p< .001)	ACHMOT	.152	1.068	.300

TABLE 6.12.2.3

Analyses of the four selected covariates (N-SCATT, CSAQ-Trait, V-SCATT, ACHMOT) plus SE multiple regressions on long term treatment effects on test anxiety

Dep. Variable (Mul. R, F, Sig. of T)	Covariate	Beta	T-Value	Sig. of T
AATD-F	SE	-.100	-.590	.563
(R=.748,	ACHMOT	-.329	-1.930	.072
F(5,16)=4.076,	CSAQ-Trait	.304	1.776	.095
p< .014)	V-SCATT	-.243	-1.384	.185
	N-SCATT	-.365	-2.099	.052
ITA	SE	-.191	-1.428	.172
(R=.852,	ACHMOT	-.153	-1.140	.271
F(5,16)=8.449,	CSAQ-Trait	.725	5.355	.001
p< .001)	V-SCATT	-.231	-1.666	.115
	N-SCATT	.071	.516	.613
TASEQ	SE	.129	.890	.387
(R=.823,	ACHMOT	.129	.884	.390
F(5,16)=6.710,	CSAQ-Trait	-.491	-3.342	.004
p< .002	V-SCATT	.421	2.799	.013
	N-SCATT	.208	1.398	.181

TABLE 6.12.2.4

Analyses of the four selected covariates (N-SCATT, CSAQ-Trait, V-SCATT, ACHMOT) plus ET multiple regressions on long term treatment effects on test anxiety

Dep. Variable (Mul. R, F, Sig. of T)	Covariate	Beta	T-Value	Sig. of T
AATD-F	ET	-.164	-.939	.362
(R=.758,	ACHMOT	-.299	-1.735	.102
F(5,16)=4.313,	CSAQ-Trait	.274	1.597	.130
p< .011)	V-SCATT	-.245	-1.417	.176
	N-SCATT	-.378	-2.198	.043
ITA	ET	-.085	-.576	.573
(R=.835,	ACHMOT	-.163	-1.123	.278
F(5,16)=7.346,	CSAQ-Trait	.709	4.894	.001
p< .001)	V-SCATT	-.232	-1.588	.132
	N-SCATT	.073	.501	.623
TASEQ	ET	.223	1.528	.146
(R=.839,	ACHMOT	.088	.610	.550
F(5,16)=7.621,	CSAQ-Trait	-.450	-3.145	.006
p< .001)	V-SCATT	.423	2.942	.010
	N-SCATT	.227	1.579	.134

Insert Table 6.12.2.5 here

Insert Table 6.12.2.6 here

Insert Table 6.12.2.7 here

Tables 6.12.2.3, 6.12.2.4, 6.12.2.5, 6.12.2.6, and 6.12.2.7 show the analyses of the regressions when in turn SE, ET, EWJ, Expectancy, and CSAQ.S were the added covariates in the MANCOVA. We can see that none of them produced a significant T-value nor significantly enhanced the size of the three multiple Rs. Nonetheless the addition of CSAQ.S into the ANCOVA produced substantial increases in the CSAQ-Trait Beta weights. This was probably due to the beta weights produced by the CSAQ.S which were of the opposite sign to the CSAQ-Trait. Given that the size of the multiple Rs was little changed these two measures were probably accounting for the same variance.

6.12.3 Long term treatment effects on test anxiety and moderator variables

Table 6.12.3.1 shows the test anxiety means at pre-test

TABLE 6.12.2.5

Analyses of the four selected covariates (N-SCATT, CSAQ-Trait, V-SCATT, ACHMOT) plus EWJ multiple regressions on long term treatment effects on test anxiety

Dep. Variable (Mul. R, F, Sig. of T)	Covariate	Beta	T-Value	Sig. of T
AATD-F	EWJ	-.084	-.364	.722
(R=.741,	ACHMOT	-.301	-1.532	.148
F(5,16)=3.410,	CSAQ-Trait	.345	1.564	.140
p< .032)	V-SCATT	-.247	-1.251	.231
	N-SCATT	-.348	-1.855	.085
ITA	EWJ	.140	.773	.452
(R=.850,	ACHMOT	-.173	-1.122	.281
F(5,16)=7.309,	CSAQ-Trait	.662	3.828	.002
p< .001)	V-SCATT	-.209	-1.347	.199
	N-SCATT	.055	.372	.716
TASEQ	EWJ	-.099	-.506	.621
(R=.824,	ACHMOT	.127	.765	.457
F(5,16)=5.929,	CSAQ-Trait	-.487	-2.617	.020
p< .004	V-SCATT	.351	2.105	.054
	N-SCATT	.215	1.359	.196

TABLE 6.12.2.6

Analyses of the four selected covariates (N-SCATT, CSAQ-Trait, V-SCATT, ACHMOT) plus Expectancy multiple regressions on long term treatment effects on test anxiety

Dep. Variable (Mul. R, F, Sig. of T)	Covariate	Beta	T-Value	Sig. of T
AATD-F	Expectancy	.115	.540	.597
(R=.747,	ACHMOT	-.380	-2.126	.049
F(5,16)=4.050,	CSAQ-Trait	.351	1.823	.087
p< .014)	V-SCATT	-.276	-1.481	.158
	N-SCATT	-.334	-1.852	.083
ITA	Expectancy	.013	.071	.944
(R=.831,	ACHMOT	-.192	-1.281	.281
F(5,16)=7.135,	CSAQ-Trait	.730	4.527	.001
p< .001)	V-SCATT	-.234	-1.499	.153
	N-SCATT	.085	.566	.579
TASEQ	Expectancy	-.005	- .026	.980
(R=.813,	ACHMOT	.154	.982	.341
F(5,16)=6.243	CSAQ-Trait	-.492	-2.919	.010
p< .002)	V-SCATT	.422	2.581	.020
	N-SCATT	.199	1.263	.225

TABLE 6.12.2.7

Analyses of the four selected covariates (N-SCATT, CSAQ-Trait, V-SCATT, ACHMOT) plus CSAQ.S multiple regressions on long term treatment effect on test anxiety

Dep. Variable (Mul. R, F, Sig. of T)	Covariate	Beta	T-Value	Sig. of T
AATD-F	CSAQ.S	-.144	-.521	.610
(R=.747,	ACHMOT	-.358	-2.115	.051
F(5,16)=4.041,	CSAQ-Trait	.413	1.524	.147
p< .015)	V-SCATT	-.237	-1.342	.198
	N-SCATT	-.306	-1.521	.148
ITA	CSAQ.S	-.159	-.697	.496
(R=.836,	ACHMOT	-.200	-1.435	.171
F(5,16)=7.446)	CSAQ-Trait	.845	3.778	.002
p< .001)	V-SCATT	-.224	-1.539	.143
	N-SCATT	.141	.848	.409
TASEQ	CSAQ.S	.394	1.780	.094
(R=.847,	ACHMOT	.182	1.347	.197
F(5,16)=8.112	CSAQ-Trait	.789	-3.636	.002
p< .001)	V-SCATT	.404	2.865	.011
	N-SCATT	.056	.348	.732

and follow-up 12 adjusted for the four significant moderator variables: ACHMOT, CSAQ-Trait, V-SCATT, N-SCATT.

Insert Table 6.12.3.1 here

Figures 6.12.3.1, 6.12.3.2, and 6.12.3.3 graphically show respectively the AATD-F, ITA, and TASEQ group means.

Insert Figure 6.12.3.1 here

Insert Figure 6.12.3.2 here

Insert Figure 6.12.3.3 here

Tables 6.12.3.2 and 6.12.3.3 show the results of the MANCOVA and ANCOVAs respectively. Again we see that as when no covariates were introduced (c.f. Tables 6.12.1.2 and 6.12.1.3) there was a highly significant effect for time at both the multivariate and the univariate level; moreover there were highly significant regressions at both levels of analysis.

TABLE 6.12.3.1

Test anxiety measures, Pre-test and Follow-up 12 group means adjusted for ACHMOT, CSAQ-Trait, V-SCATT, and N-SCATT

Group	n	Time of measurement	AATD-F	ITA	TASEQ
1	6	Pre-test	19.09	44.45	905.48
		Follow-up 12	7.21	36.41	1130.90
2	5	Pre-test	17.50	46.83	719.57
		Follow-up 12	- 2.56	30.13	1484.78
3	8	Pre-test	17.53	47.34	852.54
		Follow-up 12	2.73	37.43	1437.82
4	6	Pre-test	16.12	44.15	723.92
		Follow-up 12	6.20	35.25	1084.53

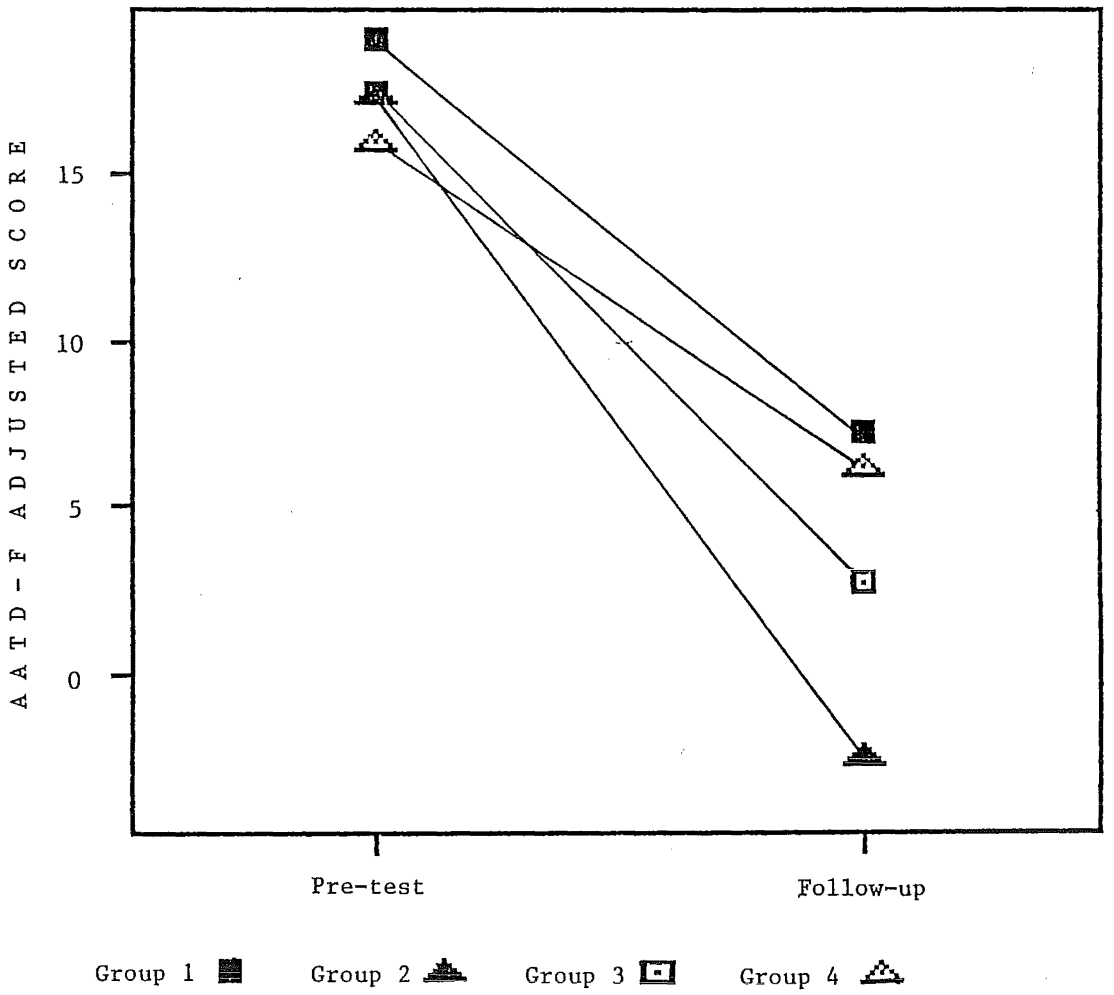
FIGURE CAPTIONS

Figure 6.12.3.1: AATD-F Group Means at Pre-test and Follow-up 12, adjusted for ACHMOT, CSAQ-Trait, V-SCATT, N-SCATT.

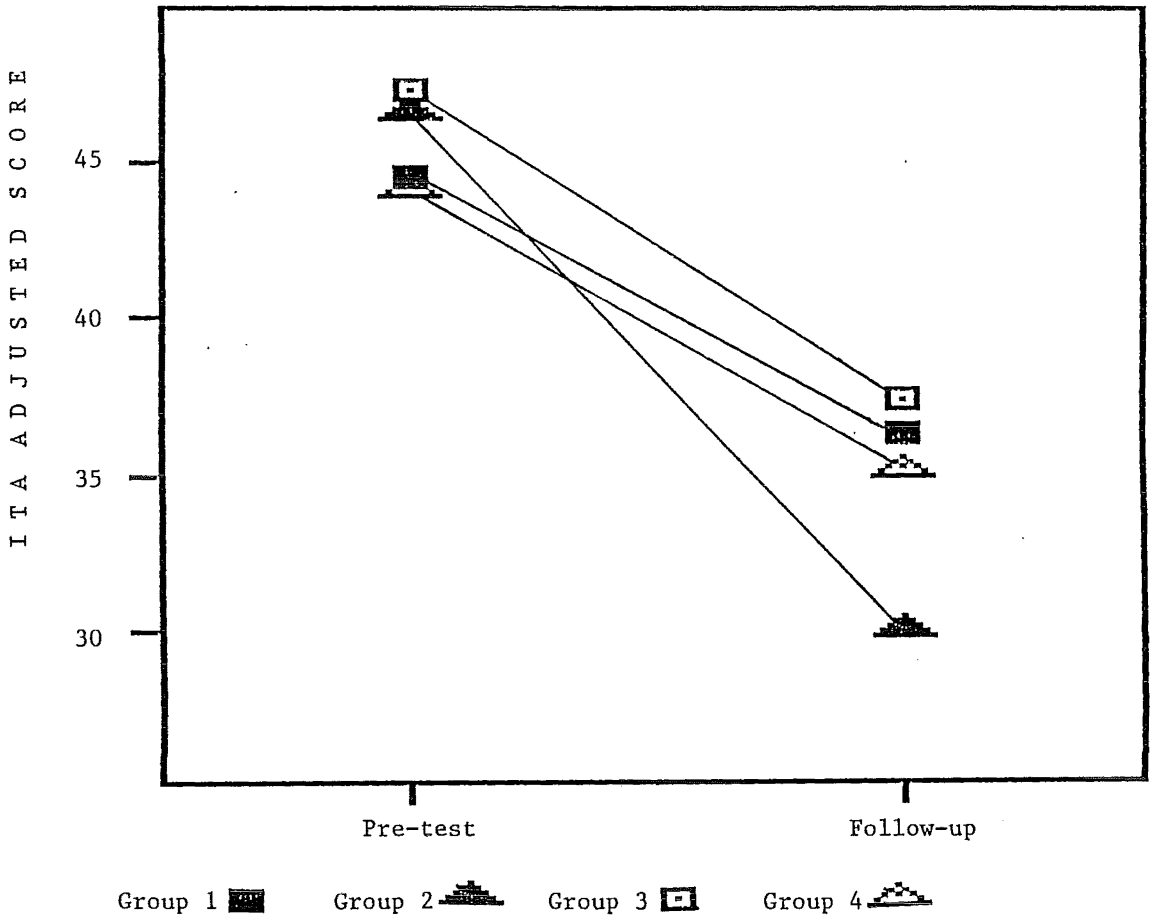
Figure 6.12.3.2: ITA Group Means at Pre-test and Follow-up 12, adjusted for ACHMOT, CSAQ-Trait, V-SCATT, N-SCATT.

Figure 6.12.3.3: TASEQ Group Means at Pre-test and Follow-up 12, adjusted for ACHMOT, CSAQ-TRAIT, V-SCATT, N-SCATT.

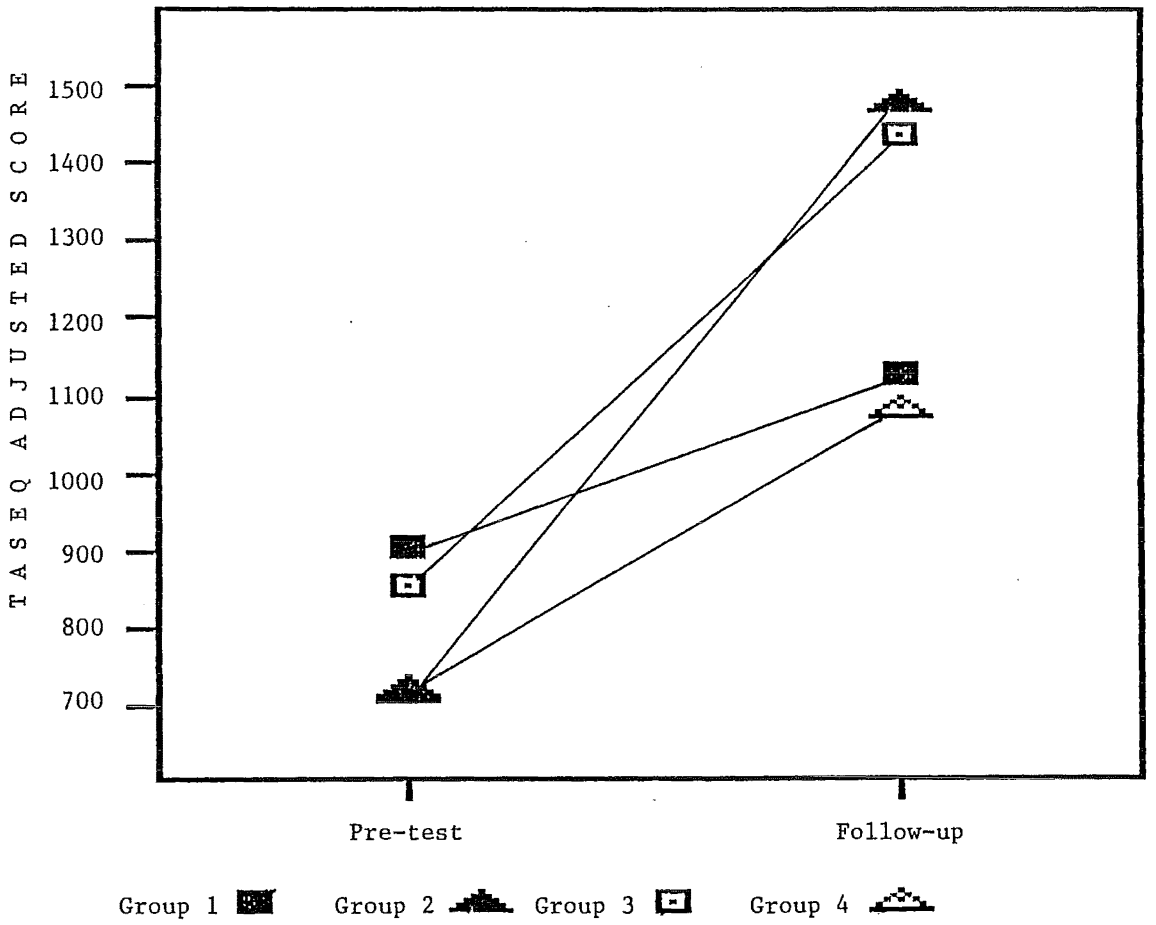
F I G U R E 6.12.3.1.



F I G U R E 6.12.3.2.



F I G U R E 6.12.3.3.



In agreement with the tables referred to above the group by time and group effects did not even approach significance. Note how the group by time significance level remained remarkably unchanged on all four analyses when the covariates were introduced. On the other hand the significance level of the group effect on the AATD-F and ITA analyses was substantially enhanced by the introduction of the four covariates: on the TASEQ analysis it was dramatically enhanced.

6.13 Covariates Correlations and Intercorrelations

In this section I am going to report the correlations of the ten covariates here studied with the three test measures at the four measurement times, and the ten covariates intercorrelations.

6.13.1 Covariates correlations with test anxiety measures

Tables 6.13.1.1, 6.13.1.2, and 6.13.1.3 show the correlations of the ten covariates here studied with the three test anxiety measures at the four measurement times.

Insert Table 6.13.1.1 here

TABLE 6.13.1.1

Covariates correlations (significance level) with AATD-F at pre-test, post-test, follow-up, and follow-up 12

Covariate	Pre-test (n=37)	post-test (n=37)	follow-up (n=36)	follow-up 12 (n=27)
E-NZ	-.40(.007)	-.35(.018)	-.40(.008)	-.30(.068)
CSAQ.S	.05(.381)	.16(.178)	.08(.323)	.20(.160)
CSAQ.Trait	.27(.054)	.16(.169)	.24(.078)	.43(.015)
ACHMOT	-.16(.176)	-.31(.035)	-.22(.101)	-.43(.013)
V-SCATT	-.33(.022)	-.33(.023)	-.24(.077)	-.35(.054)
N-SCATT	-.46(.002)	-.33(.023)	-.23(.085)	-.35(.042)
SH	-.10(.285)	-.03(.425)	-.23(.081)	-.41(.019)
SE	.15(.186)	.03(.435)	.06(.359)	-.25(.106)
ET	.07(.349)	.14(.205)	-.02(.454)	-.31(.062)
EWJ	.13(.241)	.24(.093)	-.13(.230)	.07(.371)
Expectancy	-.19(.131)	-.30(.037)	-.27(.053)	-.17(.197)

Insert Table 6.13.1.2 here

Insert Table 6.13.1.3 here

The strongest trend evident from these tables is that those covariates that were found to be significant in the previous section (ACHMOT, CSAQ-Trait, V-SCATT, N-SCATT) tended to either correlate most highly and most significantly at the first and last measurement time or to correlate significantly on all four measurement times. The other covariates tended to correlate (if at all) either at pre-test or at follow-up 12 with two exceptions: E-NZ and Expectancy.

Expectancy tended to produce sizeable correlations which were significant or approach significance with all three measures at all four times of measurement. E-NZ produced relatively large and significant correlations with AATD-F.

6.13.2 Covariates intercorrelations

TABLE 6.13.2.1 shows the ten covariates intercorrelations.

Insert Table 6.13.2.1 here

TABLE 6.13.1.2

Covariates correlations (significance level) with ITA at pre-test, post-test, follow-up and follow-up 12

Covariate	Pre-test (n = 37)	Post-test (n = 37)	Follow-up (n = 36)	Follow-up 12 (n = 27)
E-NZ	-.23(.086)	-.22(.095)	-.04(.406)	-.02(.468)
CSAQ.S	.31(.029)	.13(.214)	.09(.299)	.19(.179)
CSAQ.Trait	.48(.001)	.16(.169)	.12(.233)	.39(.026)
ACHMOT	-.02(.449)	-.21(.115)	-.31(.032)	-.36(.034)
V-SCATT	-.34(.019)	-.44(.003)	-.14(.205)	-.22(.142)
N-SCATT	-.08(.313)	-.28(.049)	-.09(.289)	-.06(.386)
SH	-.31(.031)	-.20(.114)	-.19(.136)	-.33(.051)
SE	.10(.278)	-.10(.273)	.11(.250)	-.19(.171)
ET	.03(.438)	.03(.432)	.04(.414)	-.24(.118)
EWJ	.46(.004)	-.03(.437)	.07(.345)	.35(.052)
Expectancy	-.30(.038)	-.18(.149)	-.22(.099)	-.21(.152)

TABLE 6.13.1.3

Covariates correlations (significance level) with TASEQ at pre-test, post-test, follow-up and follow-up 12

Covariate	Pre-test (n = 37)	Post-test (n = 37)	Follow-up (n = 36)	Follow-up 12 (n = 27)
E-NZ	.54(.001)	.16(.181)	-.02(.464)	.23(.119)
CSAQ.S	-.29(.041)	-.10(.285)	.13(.219)	-.13(.252)
CSAQ.Trait	-.46(.002)	-.14(.208)	-.13(.216)	-.44(.011)
ACHMOT	-.07(.348)	.22(.098)	.14(.201)	.29(.071)
V-SCATT	.35(.017)	.36(.015)	.38(.010)	.44(.012)
N-SCATT	.32(.026)	-.03(.437)	.01(.465)	.22(.139)
SH	.14(.211)	.31(.030)	.36(.014)	.54(.002)
SE	-.05(.376)	.09(.302)	.07(.338)	.17(.201)
ET	.05(.392)	.16(.165)	.14(.211)	.39(.022)
EWJ	-.40(.001)	.14(.221)	-.14(.216)	-.22(.146)
EXpectancy	.19(.131)	.46(.002)	.24(.079)	.28(.076)

TABLE 6.13.2.1

Covariates intercorrelations

	CSAQ.S	CSAQ-Trait	ACHMOT	V-SCATT	N-SCATT	SH	SE	ET	EWJ	Expectancy
E-NZ	-.36 (.016)	-.38 (.011)	-.07 (.332)	-.06 (.362)	-.42 (.005)	-.05 (.393)	-.19 (.132)	-.11 (.258)	.06 (.374)	.10 (.273)
CSAQ.S		.70 (.001)	-.15 (.185)	.12 (.241)	.03 (.426)	-.22 (.099)	-.05 (.373)	-.01 (.472)	.13 (.228)	-.23 (.082)
CSAQ-Trait			-.11 (.255)	-.02 (.457)	-.16 (.171)	-.44 (.003)	-.21 (.111)	-.29 (.042)	.30 (.045)	-.31 .(029)
ACHMOT				-.05 (.381)	-.02 (.460)	.12 (.243)	.19 (.139)	.15 (.186)	.08 (.322)	.41 (.006)
V-SCATT					.37 (.012)	.24 (.073)	-.08 (.322)	-.07 (.347)	-.24 (.089)	.23 (.083)
N-SCATT						-.15 (.185)	-.06 (.356)	-.12 (.231)	-.09 (.306)	-.17 (.157)
SH							.37 (.011)	.28 (.044)	.03 (.432)	.33 (.023)
ET								.67 (.001)	.07 (.342)	-.01 (.467)
EWJ									-.03 (.425)	-.03 (.439)
Expectancy										-.11 (.271)

Apart from covariates which were expected to intercorrelate significantly, (the scholastic ability, study skills, and general anxiety measures) there were very few significant intercorrelations.

Expectancy tended to correlate with several other covariates. This suggests that its correlations with test anxiety measures presented in the previous section and its acting as a highly significant covariate in the ANCOVA for general anxiety was a spurious effect considering that expectancy never produces a significant T-value in the covariates regressions analyses when entered concurrently with other covariates, (c.f. sections 6.3, 6.4, 6.12).

SH correlations with CSAQ-Trait and V-SCATT also explain why, although producing significant betas when entered in the ANCOVAs by itself, it failed to do so when entered together with other significant covariates.

E-NZ correlations with CSAQ-Trait, CSAQ.S and N-SCATT similarly help to explain how it was no longer a significant covariate in the long term treatment effects on test anxiety ANCOVAs, even though it correlated as reported in the previous section with the test anxiety measures, particularly with AATD-F. On this test anxiety measure the CSAQ-Trait and N-SCATT correlations' size and significance declined from pre-test to post-test and follow-up (c.f. Table 6.13.1.1) while E-NZ's remained steadily high and significant: this explains why E-NZ was a significant covariate in sections 6.3.2 and 6.3.3. When we consider only pre-test and follow-up 12 we see that N-SCATT and CSAQ-Trait correlations with test anxiety measures remained steady or were enhanced while E-NZ's decreased: this explains why E-NZ

was no longer a significant covariate in the long term effects of treatment on test anxiety (c.f. section 6.12).

Finally, of relevance to the validity of the ethnocentrism scale (E-NZ) and the EWJ test employed in this study is the fact that they did not correlate at all.

CHAPTER VII

DISCUSSION

In this chapter there will be a discussion first of the results relating to each of the thirteen sections of results reported in the previous chapter, followed by a general discussion.

7.1 Groups Before Treatment

All the analyses carried out on pre-treatment data show that there were no significant differences among the four treatment groups. All group effects were far removed from significance levels with most Fs being well below 1 and their significance being .500 or greater in most cases. Even on the academic courses taken by their respective subjects, the groups did not appear to differ.

Where data for the control group were available the relevant analysis revealed no between-group differences at pre-test.

A point to note is that, although there was a spread of scores on the E-NZ scale, the scores were actually quite low (c.f. Table 6.1.6.1).

7.2 Social Validity

The pattern of data pertaining to social validity reveals some very interesting results. First and most importantly,

the four prcedures were well received by the subjects in all four treatment groups. Expectations for improvement were high and so was the likelihood of subjects' recommendation to a friend following the first and crucial treatment session. In all probability it is during this session that subjects made up their mind as to the potential of treatment in helping them with their test anxiety. Following treatment both the therapist and each of the programs received moderately high ratings ranging between 4.5 and 5.5 on the 7-point scales.

7.2.1 Following the first treatment session

Data from the FIT questionnaires revealed that the four treatments were equally likely to be recommended to a friend. However, expectations of improvement differed among groups, with subjects from the PK treatment expecting to improve significantly more than the rest while subjects in the P group expected significantly less than the rest.

As I tried to be equally convincing with all groups and have no reason to believe that I failed in this attempt it is likely, as Wilkins (1985) has argued, that expectancy instilled by a given psychotherapy procedure is specific to that procedure. In other words, subjects evaluated for themselves to a good extent whether the therapies were going to help them in managing their test anxiety and to what degree. Of course the rationale and treatment relating to the PK group are far more credible and impressive than those received by the P group.

When expectancy was entered as the only covariate in the MANCOVA for test anxiety measures it failed to reach

significant levels (c.f. Table 6.3.3.1). Moreover, when it was entered together with the selected covariates its contribution to the total covariance failed to even go beyond the .400 significance levels on any of the three test anxiety measures (c.f. Table 6.3.3.9). More dramatic results were obtained with respect to the general anxiety analyses where expectancy was found to be a highly significant covariate when entered singly in the ANCOVA, (c.f. 6.4.3.1). However when entered together with the other significant covariates its significance failed to go beyond the .500 level (c.f. Table 6.4.3.2).

Taking into account these results and the pattern of intercorrelations reported in section 6.13.2 (where we found that expectancy, unlike other personality variables, tends to produce several significant intercorrelations), the conclusion is inescapable that expectancy instilled by psychotherapy techniques is as much a function of the client's characteristics as it is a function of the therapy.

7.2.2 Following treatment

The PRQ yielded high alpha coefficients, and it is interesting that there was a substantial coefficient increase from the post-test measurement to the follow-up measurement. This may not be easily explained. Perhaps 4 weeks after treatment the overall impression and feelings about the treatment influenced scores on each of the PRQ items as memories specific to individual items had faded, or was it the effects of sitting final exams which enhanced the PRQ alpha coefficient?

Groups did not differ on PRQ total scores nor on each of its first five items, all resulting Fs being far from significant levels and most of them above .400. Nonetheless, here was a substantial and steady decrease in the PK subjects' likelihood of recommending a similar treatment to a friend (c.f. Tables 6.2.2.6 and 6.2.2.7) even though the effect was not significant. The fact that there are relatively large standard deviations at post-test and follow-up on this measure when PK subjects are considered confirms my impression that there were large disparities in the way the PK treatment was received. Much larger standard deviations for this group were also evident on the PRQ scores (c.f. Table 6.2.2.1).

Interestingly enough, PK subjects' ratings on the first five PRQ items tended to about as high as the other three groups but substantially lower on FRIEND. Now, the first five items relate to factual and experiential aspects of the program while the likelihood of recommending the treatment they just received to a friend calls on subjects' affective reactions to that treatment. In other words, notwithstanding the fact that PK subjects' affective reactions to their treatment were not as favourable as that of subjects from the other three groups, this did not seem to affect their ratings of the therapist and their program, which were about the same as the other subjects. The PK superiority at decreasing test anxiety and general anxiety as well as increasing academic performance evident in sections 6.3, 6.4, and 6.6 cannot be attributed to therapist bias.

Comments on their respective treatment probed on the PRQ indicated that some PK subjects resented the additional work and time demanded by the program at a time when they were very

busy preparing for exams. Moreover, the PK program covered twice as much work as the others the, pace being far more rapid during those sessions, probably at the expense of enjoyment.

The most frequent suggestion on how to improve the program was to have it at the beginning of the year so that its benefits could be reaped over the entire academic period. The next most frequently made suggestion was to make it available to all students. I don't believe the intention behind those comments was to make me feel good, but they did.

7.3 Treatments and Test Anxiety

The strongest evidence coming from section 6.3 is that all treatment groups, including Placebo, reported large decreases in test anxiety across all three measures. Although no significant between-group difference was apparent when the moderating effects of personality variables were left out, once the salient (i.e., the actual and nor spurious) covariates are introduced in the analyses we see that we get a significant group effect on the AATD-F and on the ITA and a significant group by time effect on the TASEQ. In each of these cases PK is the group that has decreased test anxiety more than the other three groups.

The introduction of the covariates in the various analyses seems to bring dramatic enhancements on the significance of the group effect but the significance of the group by time effect appears little affected if at all. Prior to the introduction of the salient covariates the significance levels for the group effect were .547, .275 and .521 for the AATD-F,

ITA, and TASEQ respectively, but when the covariates are accounted for they become .048, .022, and .331. The corresponding significance levels for the group by time effect are .553, .418, and .052 before the salient covariates are introduced and .561, .414, and .042 following their introduction. The corresponding changes on the multivariate analyses were .444 to .068 on the group effect and .309 to .311 on the group by time effect. In sections 6.4, 6.11, and 6.12 only two measurements on either test anxiety or general anxiety were considered; ANOVAs were also carried out first without salient covariates and then by introducing such covariates. We find here that the significance of the group effect was greatly enhanced and might be dramatically enhanced whereas the group by time effect remained remarkably unaffected by the introduction of covariates. This could be the result of structural problems within the ANCOVA.

The significance of the time effect, like the group by time effect, was slightly changed by the introduction of the significant covariates on the analyses where three measurements were involved but remained remarkably unchanged when only two measurements are considered.

Apart from the superiority of the PK group, the order of the treatments' effectiveness in reducing test anxiety was not as had been predicted (c.f. section 2.16). Even though no other group (except PK) appeared significantly different from the others, on both the AATD-F and the ITA Placebo subjects were the second best group while they were the third best on the TASEQ. This applies whether the effect of the covariates is taken into account or not. CATH group was the second best on both the TASEQ ANOVA and ANCOVA but came last on the AATD-F

and ITA ANOVAs and second on these respective ANCOVAs. The SES group tended to be the least effective treatment.

On the whole, the PK group was significantly superior to the other groups, and even though few othe orthogonal contrasts were significant (c.f. Table 6.3.4.4) the effectiveness of the three groups appeared to follow this order: Placebo second, CATH third, and SES fourth.

7.3.1 Salient moderating variables

The salient covariates' (E-NZ, CSAQ-Trait, V-SCATT, and ACHMOT) pattern of beta weights and their significance levels reveal an interesting story (c.f. Table 6.3.2.10). First, a different set of covariates seemed to affect the decreases on the three test anxiety measures which suggests that they were qualitatively different from each other. However, from an examination of the above table as well as sections 6.4 and 6.9 it appears that the AATD-F and TASEQ are more similar scales while the ITA differs more. As discussed in section 3.1.3 the ITA is a measure of state test anxiety while the AATD-F is a measure of trait test anxiety. Because the TASEQ had instructions which probed subjects' test anxiety self-efficacy feelings relating to a hypothetical examination or test, it acted also as a measure of trait test anxiety.

Nonetheless, the direction of the covariates' effect, either facilitating or inhibiting test anxiety decrease, was consistent across the three measures. It is rather in its magnitude and its statistical significance that it varies.

A negative beta weight indicates that the particular

covariate facilitates test anxiety decrease and vice versa when we consider the AATD-F and the ITA, while the reverse is true with respect to the TASEQ.

An examination of the table noted above indicates that V-SCATT was the only covariate to consistently and significantly affect changes on all three test anxiety scales.

The E-NZ produced the highest and most significant beta weight among all those reported on Table 6.3.2.10, specifically on the AATD-F regression. Nonetheless, it approached significance on the TASEQ regression and was clearly not significant on the ITA. Interestingly enough, I had predicted that rigidity, the construct intended to be measured by E-NZ, would inhibit the course of treatment while in fact it appears to have facilitated such a course.

General anxiety (CSAQ-Trait) seems to have been the only covariate to significantly inhibit the course of treatment. It produced a significant beta weight on the ITA however. While this covariate beta weight failed to approach significance on the TASEQ regression, it is clear that it had virtually no effect on the AATD-F decreases, judging by the extremely low beta weight (remote from even approaching significance) it produced on this regression.

Achievement motivation (ACHMOT) produced a similar pattern as the general anxiety one but in the opposite direction and on different test anxiety scales. It produced a significant beta weight on the AATD-F regression, one that was not significant on the ITA regression, and one that was remote from significance on the TASEQ.

Finally, we can see from Table 6.3.2.10 that the four salient covariates produced quite large multiple regression

coefficients on all three test anxiety measures, and that these were of similar magnitude.

7.3.2 Worry versus emotionality

Liebert and Morris's (1967) formulation of test anxiety would predict that PK and CATH subjects would show decreases in worry but not emotionality following treatment or at least that worry should decrease more than emotionality. The analysis carried out in section 6.3.5, however, shows that these subjects tended to decrease more on emotionality than on worry.

The results reported in section 6.3.6 show that there was no significant difference in the PK and CATH subjects' tendency to increase on the worry as opposed to the emotionality scale of the TASEQ, which appear to be valid and reliable subscales.

These findings corroborate previous similar experiments cited in section 1.3. On the whole these findings suggest that the distinction between worry and emotionalility is not as clear cut as Liebert and Morris' theory purports.

7.4 Treatments and General Anxiety

Although the way the general anxiety index was obtained is no doubt going to provide critical ammunition to the cynic, nonetheless I believe this may be a better way to measure changes in general anxiety than administering the same scale before and after treatment. This is because it prevents practice effects, which are always present in multiple administrations of the same index, from confounding post-test scores. Of course this is not a problem with the test anxiety measures (administered four times) because countless studies have shown that non-treated subjects do not show significant changes in self-reported test anxiety measured by the same test anxiety measure administered several times.

Finally, correlations between general anxiety measures and test anxiety measures (c.f. section 6.9.2) clearly suggest that general anxiety had changed from pre-treatment levels.

7.4.1 Which treatment was more effective?

The results reported in sections 6.4.1 and 6.4.2 show that PK subjects decreased on general anxiety more than any of the other groups. CATH subjects also decreased significantly on general anxiety although without producing a significant orthogonal contrast. Placebo subjects also showed decreases on this measure but these were clearly not significant; on the other hand SES subjects appear not to have changed on general anxiety at all.

It is interesting to note from Figure 6.4.1.1 that there seem to be large differences between the four slopes

indicating changes on general anxiety. Further, there are five crossovers, yet the group by time effect is just within significant levels ($p<.047$)

Finally, treatment effects on general anxiety seem to reflect more closely the psychological power basis of the four treatments.

7.4.2 Effect of moderator variables

It's interesting that as many as six covariates (ITA, AATD-F, TASEQ, SH, E-NZ, and Expectancy) formed mostly large and highly significant beta weights in the ANCOVAs for general anxiety, yet that when entered simultaneously in the analysis only two (ITA and SH) remained significant (c.f. Tables 6.4.3.1 and 6.4.3.2). The simulataneous investigation of the covariates which individually produced significant beta weights has produced some rather dramatic changes to both the size and significance of these variables effects' on the course of treatment.

Of particular interest is Expectancy, which when entered on its own appeared to produce a substantial and highly significant facilitative effect on the course of treatment in reducing general anxiety, but when entered simultaneously with the other significant correlates produced a very small beta weight which was far from even approaching significance.

A similar state of affairs comes about when we examine AATD-F and TASEQ; interestingly enough, although these two measures and the ITA presumably measured the same construct and all three produce relatively large and highly significant betas when entered individually in the ANCOVA for general

anxiety, when entered simultaneously there was strong and clear evidence that there is something distinctive about the nature of ITA (also shared by AATD-F and the TASEQ) which inhibits the course of treatment from decreasing general anxiety. On the other hand, AATD-F and TASEQ beta weights became small or very small and remote from even approaching significant levels.

The change in the E-NZ beta, although not as drastic as that of Expectancy, AATD-F and TASEQ, followed the same pattern. SH, although with a much reduced beta in the multiple covariates ANCOVA, was still acting as a treatment facilitator which was well within significance levels.

When AATD-F, TASEQ, E-NZ and Expectancy were excluded from the above ANCOVA, ITA beta weight and its significance were substantially enhanced but the SH equivalent was hardly changed. This suggests that there is something in the nature of the four previously excluded covariates which is shared by ITA while SH facilitating effects on the course of treatment at reducing general anxiety seem to have a highly distinctive nature.

The introduction of the salient covariates (ITA and SH) into the ANOVA for general anxiety brought absolutely no change on the significance of the time nor the group by time effect, yet there was an enhancement of the significance of the group effect.

7.5 Treatments and Study Skills

The results reported in section 6.5 clearly show that the two groups which received SES gained the predicted benefits

and showed significant improvements on study skills measures SH, SE, and ET. Although the multivariate analyses showed a clear overall improvement, the univariate analyses pointed out that there were significant improvements on the SH and ET measures but the improvements on SE were approaching significance with respect to the SES group and were virtually significant with respect to the PK group.

Interestingly enough, P subjects showed a strong improvement on the study skills measures but without reaching the significance zones.

7.6 Treatments and Academic Ability

In spite of the fact that data from many of the subjects were not available, significant effects were obtained on academic performance improvements particularly with respect to the more valid of the two measures (FGA) and particularly in the more crucial range of ability (between the 15th and 85th percentile of the N-SCATT distribution, c.f. section 6.8)

7.6.1 Group differences in FGA improvements

We saw in section 6.6.1 that the superiority of the PK group did not reach significant levels even when the control subjects were included in the analyses. However, in section 5.6.2 it became apparent that the superiority of the PK treatment approached significance in the SCATT middle range and was not significant in the V-SCATT middle range, but that in the N-SCATT middle range it was clearly significant. Moreover, results reported in section 6.8 showed that, at

least for the sample of students here investigated, the N-SCATT broad middle range was crucial as far as the negative relationship between test anxiety and academic performance was concerned.

Not only was the relationship between these two variables stronger in this range when compared with the other ranges of ability but larger decreases in test anxiety were associated with larger increases in FGA more strongly in the N-SCATT broad middle range. Of course, this suggests that the importance of the SCATT middle range in the relationship between test anxiety and academic performance (c.f. section 1.2.2.1) may be due to the N-SCATT subtest while the V-SCATT subtest (which is added to N-SCATT to give the total SCATT score) acts to dampen the importance of N-SCATT in this issue.

Unfortunately none of the studies that have investigated the effects of ability on the relationship between anxiety or test anxiety and academic performance (Gaudry & Fitzgerald, 1971; Paul & Eriksen, 1964; Spielberger, 1962; Spielberger & Katzenmaier, 1959) have attempted to tease out the aspect of ability which is the crucial one.

In spite of the fact that more than a third of treatment subjects where not at university in 1982 and did not have pre-test FGA scores, and in spite of the fact that of the remaining 23 only 12 had an N-SCATT score in the designed range, the ANOVA showed a significant time effect (the same as when all 23 subjects were considered) and a significant group by time effect. I hardly need to point out that obtaining significant results with such small N is truly remarkable. Moreover, it indicates two important points. Firstly, that highly test anxious subjects with an N-SCATT score falling in

the broad middle range of this distribution are those who benefit most from test anxiety treatment efforts to improve academic performance, and secondly, in this crucial range the PK treatment superiority in achieving this objective is demonstrated to be significant. It is important to note that subjects with an N-SCATT score falling at either extreme of this distribution would not be expected to improve their academic performance as there appears to be no negative relationship between test anxiety and academic performance for these individuals. Therefore if we include these subjects in an ANOVA trying to detect differences between test anxiety treatments in improving academic performance, we add systematic error in such analysis, thereby substantially reducing its power at identifying possible treatment differences. Actually, the greater the number of subjects described above in such ANOVA the more impaired its power at achieving what it is intended to achieve: identification of true between-group differences along the intended dimension.

Interestingly enough, it is in this crucial N-SCATT range that the effectiveness of the four treatments in improving academic performance follows more closely the predicted order (c.f. section 2.16) with PK being first and CATH second.

7.6.2 Group differences in TEA changes

In support of previous findings (Gaudry & Bradshaw, 1970) and formulations (S. Sarason et al., 1960), subjects in this sample obtained higher marks in exams held during the year than in final exams. Nonetheless, in this institution the custom in many courses of including easier questions in

internal exams than in final exams (c.f. section 4.3.3) made the above decrease more marked than expected.

We see from Table 6.6.3.2 that there was no between group difference on this measure. However an examination of section 6.8 indicates why. Even though there was negative correlation between TEA and test anxiety on pre-test scores and again this relationship was much more pronounced in the N-SCATT range, when we consider changes in test anxiety scores and changes on TEA scores we see that the relationship is the opposite of what would be expected: longer decreases in test anxiety were associated with longer decreases on TEA. Given this state of affairs, although the SES component of the PK treatment would work towards reducing the decrease in TEA scores at post-test its greater effectiveness at decreasing test anxiety would work in the opposite direction. Henceforth even though PK subjects' decrease on this measure was limited and it achieved substantially higher post-test TEA scores than the other groups, no significant between-group difference was apparent.

7.6.3 General anxiety and academic performance

As we saw in section 6.6.5 subjects low on general anxiety and subjects high on this measure did not differ in improvements on FGA nor in decreases in TEA. Actually there was a tendency for the high general anxiety subjects to benefit more from treatment along this dimension than their low counterparts. These results are not in agreement with McMillan and Osterhouse's (1972) finding that low general anxiety subjects improve on academic performance more than high general anxiety subjects in their test anxiety treatment

study.

There are two possible reasons that may explain the difference found in this study when compared with McMillan and Osterhouse (1972). The earlier study employed a different general anxiety measure (the MAS) and a different form of treatment: systematic desensitization. At this point we cannot establish whether it was the difference in the treatments or the difference in the general anxiety measures or some other reason that caused such different findings.

7.7 Treatments and Ability Tests Measures

The analyses reported in section 6.7 clearly point out that there was no improvement in the ability tests employed in this study, nor was there any between-group difference. All the effects were far removed from significance levels, producing Fs which in no respect reach unity.

As with the TEA measure, although test anxiety and ability measures were negatively related at post-test (c.f. section 6.8), larger test anxiety decreases were associated with larger decreases in ability tests performance. Therefore the test anxiety treatments designed to reduce test anxiety would not be likely to increase performance on the above ability tests. In the light of this finding we really have to question how previous test anxiety treatment investigations (c.f. Chapter II) produced improvements on ability test measure. One difference between this and past research is that a much longer and comprehensive test was employed in this study. Moreover, the post-treatment test was carried out 1

year after treatment. So there is a possibility that the subjects in this study improved on these measures in the short term but that this effect withered with time.

I was surprised to find no improvement on these ability measures , as subjects tended to finish the test at the second testing, while on the first occasion they tended to answer only about three quarters of the items. I reasoned that on the second occasion subjects were by then no longer overly concerned at being negatively evaluated (i.e., much lower on test anxiety) and were not trying so hard at what is a very demanding test in terms of cognitive-intellectual effort.

7.8 Test Anxiety and Performance Correlations

As pointed out in section 6.8, because the number of subjects is small, individual correlations between test anxiety and performance obtained in this study vary markedly; it is the trends in these correlations that are of value. Moreover, given the restricted range of scores on test anxiety measures before treatment (a criterion for subject selection was an AAT- score falling in the top 13% of the distribution of scores on this measure) it is quite remarkable that so many such correlations were of about the same size as those found in correlation research reported in the literature (c.f. section 1.2.2). Anastasi (1968, p.92-94) has explained how no correlation may be found between two variables if a sample with a restricted range of scores is chosen, whereas the true correlation between these two variables is very high if a heterogeneous sample is selected.

The results from these correlations tie in extremely well with the results reported in sections 6.6 and 6.7 and the main hypotheses of this study (i.e., that PK is a superior treatment). As already pointed out, it is when we consider the FGA measure that we obtain negative correlations with test anxiety before treatment and, most importantly, that we obtain negative correlations between test anxiety changes and FGA changes. The fact that this relationship is much stronger in the broad middle range of N-SCATT points out the importance of this ability measure range in the relationship between test anxiety and performance. Again, as has been pointed out, although at pre-test the ability tests measures and TEA correlated negatively with test anxiety, test anxiety decreases were associated with decreases on these measures. Therefore we cannot expect test anxiety treatment to improve performance measured by these indexes. It is interesting that positive correlation between changes in test anxiety and on the ability measures was strongest in the N-SCATT broad middle range where the inverse relationship between test anxiety changes and FGA changes was also found to be strongest (c.f. Table 6.8.2.2).

The positive correlation between test anxiety decreases and TEA decreases is in accord with the results obtained by Gaudry and Bradshaw (1970).

Correlations between test anxiety at the 12-month follow-up and FGA and ability measures at the same follow-up (c.f. Table 6.8.1.2) tended to disappear. Nonetheless, sizable correlations were obtained in the N-SCATT range (again the magic range) but, unlike the correlations between these variables at their pre-test levels (c.f. Table 6.8.1.1), they

were positive. This indicates that overly large test anxiety decrements following treatment may be detrimental to students' academic performance and ability tests performance, particularly for those students in the middle range of the N-SCATT distribution. We also have to bear in mind that as the spread of test anxiety scores was a lot greater at the 12-month follow-up, this reduces the problem of the validity of these correlations resulting from too restricted a range of scores on one measure (c.f. Anastasi, 1968). The spread of FGA and ability tests scores remained stable across the various measurements.

With reference to FGA, correlations between test anxiety changes and academic performance changes appear to increase with a longer interval to reach sizes well above those usually reported in the literature between test anxiety and academic performance at one point in time (i.e., snapshot correlations).

In short this validates the note of caution made previously that test anxiety researchers ought to look at the correlation between test anxiety changes and performance changes (c.f. section 2.3.2) when evaluating the effectiveness of test anxiety treatment in improving academic and/or ability tests performance.

Finally, it is evident from Tables 6.6.2.2 and 6.6.3.2 that there is no significant time effect on either the FGA or the TEA when the middle range of V-SCATT was considered but there was a significant such effect on both these measures when the N-SCATT was considered. Somehow it is in this latter range that results (FGA increases and TEA decreases) are in accord with past research. Now, as past research relevant to

this point has investigated larger and more heterogeneous groups, even a small effect would yield significant results. The point I am trying to make is: it appears that individuals with a N-SCATT score falling in the broad middle range of the distribution of scores on this measure have their academic performance more strongly affected by test anxiety.

7.9 THE TASEQ

THe data and analyses reported in section 6.9 strongly suggest that the TASEQ is a valid and reliable measure of test anxiety self-efficacy or SEMA which is tantamount to the converse of test anxiety, as has been argued in section 3.2.1. It correlates both with the ITA and the AATD-F, producing in every instance higher correlations with the latter measure. This result in turn validates the TASEQ which, if it bears instructions relating to the self-efficacy of a hypothetical exam as has been done at pre-test, post-test, follow-up and follow-up 12, like the AATD-F it is a measure of trait test anxiety. Moreover, the behaviour of the salient covariates reported in section 6.3 show the TASEQ to be qualitatively closer to the AATD-F than the ITA. TASEQ's pattern and the sizes of correlations with general anxiety measures are very similar to the other two test anxiety measures here employed (c.f. Table 6.9.2.1). Of great importance is that, even alongside the AATD-F, which is one of the best anxiety measures (c.f. section 3.1.4), the TASEQ correlates more strongly and more significantly with subjects' direct reports of test anxiety improvements as a result of treatment and with changes in such direct reports (c.f. Table 6.9.2.2)

This measure yielded some extremely high internal consistency coefficients, which is a testimony of both good validity and good reliability. It is interesting that such coefficients increased considerably from what was an already high level of internal consistency between the pre-test and the post-test measurement, and remained essentially the same as the latter measurement at follow-up. I reasoned that it was the subjects' greater familiarity with the instrument and, also importantly, their enhanced ability at evaluating their own self-efficacy as a result of treatment and practice at doing so which produced such impressive alpha reliability enhancement.

7.10 Self-Efficacy ...

As predicted in section 3.2.2.3 SEMA feelings during exams as measured by the TASEQ accurately predicted subjects experienced test anxiety as measured by the P-TASEQ. From Table 6.10.1.1 and Figure 6.10.1.1 we can see that TASEQ group means closely matched P-TASEQ group means. Moreover these two measures correlated highly at the first exam measurement and very highly at the second. Presumably as subjects became more familiar with these questionnaires the validity of their scores increased and so did the correlation between these two set of scores.

These results strongly support the validity of the SEMA-SESC distinction in self-efficacy theory. The fact that the two exams were real-life final exams is all the more reason to believe that these results and the theory they are

intended to test are valid. Moreover, the high levels of self-efficacy feelings related to these real exams is a strong test showing the validity of the treatment's effectiveness in decreasing test anxiety. It is also obvious from examining Figure 6.10.1.1 and Table 6.10.1.1 that SEMA feelings measured by the TASEQ are as accurate when they refer to a hypothetical exam as they are when they refer to a real-life exam.

7.10.1 ... And the P-TASEQ

The P-TASEQ correlated either highly or very highly with the TASEQ while the ITA-TASEQ correlations at the first and last exam reached only moderate levels. Because the ITA is an established state test anxiety index, the P-TASEQ correlations with this measure (moderate in the first and moderately high in the last exam) provide evidence supporting the construct validity of the P-TASEQ as a state test anxiety measure. Moreover, while the test-retest reliability of the ITA relating to these measurements was .62 ($p < .001$) the corresponding P-TASEQ was .83 ($p < .001$).

On the whole, considering (1) the way the P-TASEQ was constructed (ie., strong affinity to the TASEQ), (2) the high correlations between these two self-efficacy measures, (3) its moderate to moderately high correlations with the ITA, and (4) its high test-retest reliability, we can be confident that the P-TASEQ is a good test anxiety measure in terms of validity and reliability.

7.11 One Year Later

In this section I am going to discuss the results of the 12-month follow-up. The test anxiety measures, exam skills measures, and academic performance will be considered in turn.

7.11.1 Test anxiety

Overall the test anxiety decrements achieved by the four treatments were maintained. There was no significant time effect either at the multivariate or at the univariate level on all three measures. PK subjects decreased significantly on the TASEQ but there was no other significant between group difference. The advantage of the PK group over the others was entirely lost and there was a tendency in this group to report more test anxiety when compared with the 4-week follow-up although this reacted significance only on the TASEQ.

Some interesting changes occurred with the introduction of the four salient covariates in the analyses. The regression in the MANCOVA was still significant although its significance was not as high. However, a look at the analyses of the ANCOVA regressions (c.f. Table 6.11.2.3) reveals that the importance of the covariates had changed when compared to the results obtained at post-test and at the 4-week follow-up (c.f. Table 6.3.3.10). E-NZ beta weights all decreased and none of them reached significant levels. V-SCATT beta weights all decreased across all three measures and, whereas before (i.e., in Table 6.3.3.10) it produced relatively high beta weights (all significant) on the three test anxiety measures, it now produced only one significant such beta, on the ANCOVA

for TASEQ. On the other hand general anxiety (CSAQ-Trait) increased all three of its beta weights and all were at significant or near significant levels. ACHMOT increased its beta weights on the ITA and on the TASEQ but decreased slightly on the AATD-F. Nonetheless, ACHMOT appeared to have become a far more significant covariate overall.

The size of the multiple R increased substantially in the TASEQ ANCOVA, to a lesser degree on the ITA, and increased slightly on the AATD-F.

7.11.2 Exam skills

The results reported in section 6.11.3 show that there was no significant change at the 12-month follow-up on the two exam skills measures. We can conclude that the improvements obtained by the subjects who received SES training (PK and SES) on these measures were maintained.

7.11.3 Academic performance

The results reported in section 6.11.4 show that academic (FGA) improvements reported at post-test were maintained. Moreover, when the ability ranges are taken into account, as has been done in section 6.6, the superiority of PK subjects over the other treatment groups reaches significance levels only in the N-SCATT middle range. Moreover, in those ability ranges where the superiority of the PK treatment approached significance after treatment, the maintenance of such

superiority at follow-up 12 also approaches significance. In sum, these results point out not only that improvements in FGA were maintained but also the validity of the importance of taking into account ability when evaluating a test anxiety treatment's effectiveness in improving academic performance. Such validity is enhanced by the fact that many subjects who were enrolled in 1982 (i.e., the year before treatment took place) were not enrolled in 1984 (the year following treatment). In other words the analyses reported in section 6.6.2 involved a largely different set of subjects compared to the analyses carried out in section 6.11.4. Finally, the type of effect on which the PK treatment showed its superiority (the group by time effect in section 6.6.2 and the group effect in section 6.11.4), strongly indicates that (1) this treatment enhances subjects' academic performance more than the other treatments and (2) such enhancement is maintained over at least a 12 month period.

7.12 Long Term Effects of Test Anxiety Treatment

The results reported in section 6.12 suggest that ,in the long run, treatment differences decrease in importance while personality factors increase in importance, explaining reduced levels of test anxiety. Between-group differences were far from significant in the analyses reported in this section; on the other hand there were large increases, explained by personality factors, in the test anxiety measures' variance, particularly on the ITA and the TASEQ ANCOVAs.

The pattern of the significant individual personality

variables shows dramatic changes when compared to the analyses carried out in section 6.3. Most notable is the great increase in the general anxiety beta weights, which more than doubled on the ITA and the TASEQ and nearly quadrupled on the AATD-F (c.f. Tables 6.3.2.10 and 6.12.2.2) while the E-NZ was no longer significant. ACHMOT betas and their significance remained largely the same while V-SCATT's remained stable on the TASEQ but decreased quite substantially on the AATD-F and the ITA.

An interesting development took place when the CSAQ.S was introduced in the ANCOVA: CSAQ-Trait betas increased greatly. However, the overall variance explained by the covariates did not vary substantially. These two general anxiety measures produce betas of opposite signs and it appears that they were essentially explaining the same unique variance in the light of the fact that the other beta weights were largely unchanged. In sum, the substantial increase on the CSAQ-Trait beta weights was an artifact produced by the introduction of CSAQ.S

7.13 Covariates' Correlations and Intercorrelations

The results reported in section 6.13 throw light on the nature of the effects of the covariates here investigated on test anxiety. Although there appear to be more individually significant covariates than those identified in previous sections, they must be explaining a similar variance given that, when entered concurrently in the analyses, few of them remain significant.

As previously pointed out, expectancy correlates with the test anxiety
measures but it also correlates with the other covariates. Considering that this variable never produced a significant beta when entered in the ANCOVAs concurrently with other covariates we can be confident that expectancy correlations with test anxiety are spurious. It appears that differential degrees in expectancy instilled by treatments are more a function of a person's characteristics than a function of treatment differences.

This study found that ethnocentrism and performance on the EWJ test yielded no correlation at all. Although this result casts further doubt on the existence of rigidity as a construct, it bears validity for these modified scales, as both Applezwig (1954) and Goodstein (1953) found the same (i.e., no correlation at all) using the original ethnocentrism scale and the timed EWJ test (c.f. section 4.6).

7.14 General Discussion

The PK group decreased most on test anxiety and general anxiety, improved on study skills, and improved more than the other treatment groups on academic performance: with these results the main hypothesis of this project was confirmed. Moreover, we can be confident that the superiority of the PK treatment was not due to therapist bias. On the contrary, the timing of the treatment worked against the effectiveness, particularly of the PK treatment (and also against the effectiveness of the SES treatment).

CATH and SES groups also reported significant decreases in

test anxiety. The former group decreased on general anxiety and increased on academic performance, although not as much as the PK group, while the latter increased on study skills. So it appears that the superiority of the PK treatment is due to the combined contribution of both its components and that neither component is as effective as the combination of the two. This result is nicely in accordance with predictions made in section 2.16. Moreover, my impressions that a 40-minute or a 1-hour videotape on study study skills as employed by McKordick et al. (1979, 1981) is a potentially ineffective procedure (c.f. section 2.5) were justified.

These researchers employed respectively a 10- and 14-hour treatment which involved cognitive-attentional training plus a 40-minute (1979) or a 1-hour (1981) study skills videotape and did not obtain improvements on academic ability. In addition, unlike the results of this study, such short study skills procedures employed by McKordick et al. did not produce test anxiety decreases.

I believe that an effective study skills training program not only enables subjects to learn new and better study methods, it also motivates them to put such methods into practice. This is where the advantage of this SES program over that of McKordick et al. lies.

Unexpectedly, the P group reported substantial and significant reductions in test anxiety. Although the orthogonal contrasts failed to reach significant levels, P appeared to be superior than either the CATH or the SES groups. Moreover, again without reaching significant levels, P subjects decreased on general anxiety and increased on study skills. The reason for this unexpected result, I believe, is

that this procedure was far from being inert as initially intended. All subjects in this treatment appeared to achieve a deep state of relaxation during each of the seven sessions. Instructions such as "sit comfortably on chairs or comfortably lie on the floor", "think of nothing", "let intruding thoughts run their course", and "concentrate on the feelings in the various parts of the body (hands, arms, face)" clearly resemble those of meditation techniques (c.f. Cormier and Cormier, 1979; Ch 20). Furthermore, past experience in relaxing friends strongly suggested that this therapist's voice is very suggestive and his ways of relaxing are very effective.

Subjects in this group often pointed out the suggestive power of the therapist's voice and all gave clear indications that they were achieving a deep state of relaxation during all sessions. For this reason the P group has been referred to as a treatment group, not as a placebo.

The fact that Holroyd (1976) found this form of placebo treatment to be as effective as other test anxiety treatments supports this point. At this point I wish to remind the reader that Allen's (1971, 1973) placebos, when viewed in the light of Wine's (1971, 1982) theory, clearly become treatments and not placebos (c.f. sections 2.7 and 2.8).

Another cheerful outcome of this study is the fact that treatment effects were largely maintained after 1 year. Although the superiority of the PK treatment on the test anxiety decrement aspects was lost, its superiority at improving academic performance was maintained. Nonetheless, there is strong evidence which suggests that although test anxiety decrements were maintained, the degree of such

maintained anxiety decrements is more dependent on the subject's characteristics than on which treatment the subject received.

7.14.1 Moderator variables

Another important outcome from this study is the finding that personality factors are important determinants of test anxiety treatment. Once this fact is established we see that we need to control for such factors in the evaluation of differential treatment effectiveness. To this end the analyses here employed (MANCOVA and ANCOVA) appear well suited, in that they both identify and control for such covariates. Still, as has been argued in section 7.3, these statistical techniques are not without fault, in that although they appear to be far more powerful than MANOVA and ANCOVA in identifying between treatment differences on the group effect, there appears to be no difference in the power of these analyses in identifying differences on the group by time effect, particularly when only a pre-test and a post-test measurement is considered.

The role of the personality factors in test anxiety treatment is far from straightforward: whether a particular personality factor acts as a significant covariate depends not only on which measure of test anxiety we are considering but also on whether the influence of other personality factors is considered simultaneously.

In addition, we need to include the time dimension in the evaluation of whether a personality factor significantly affects the course of treatment and the extent of such effect.

One ray of simplicity in such a complex pattern is the fact that personality factors' direction of effect, either facilitating or inhibiting the course of treatment, does not change even though it may dramatically increase or decrease or disappear entirely with time or across different measures.

General anxiety is the only personality factor here investigated to significantly inhibit the course of test anxiety treatment; its effect being low in the short term and becoming the most influential personality characteristic in the long term. This result is in disagreement with McMillan and Osterhouse (1972) who found general anxiety to have no effect in the course of test anxiety treatment in decreasing test anxiety. Moreover, this is contrary to the findings of Mitchell and Ingham (1970) who found general anxiety to have a facilitating effect on the course of test anxiety treatment in the short term which disappeared at their 14-week follow-up . It must be pointed out that such differences in results do not imply inconsistency, as both these studies employed a different therapy (desensitization) and different general anxiety measures from the present study.

Verbal ability seems a powerful facilitator of test anxiety treatment in the short run, while its importance is much decreased in the long term. Achievement motivation also facilitates test anxiety treatment but, unlike verbal ability, its importance is low in the short term, increasing substantially in the long term. The facilitating effect of achievement motivation was as predicted in section 3.6.3. Numerical ability seems to have a facilitating effect only on the long term maintenance of test anxiety decrements.

Although E-NZ was intended to measure mental rigidity, I

have reservations that it did in fact measure this doubtful personality construct. I am inclined to believe that it measured ethnocentrism or authoritarianism or acquiescence to F-scale type items, as Heaven (1983) would put it.

The basis of this belief lies in the fact that ethnocentrism (will do) has had a strong treatment facilitating effect in the short term which was lost in the long term. Now, authoritarians tend to be submissive and uncritical towards authority as well as being conventional (c.f. R. Brown, 1965; Sanford, 1973). These dispositions would make those subjects who are more authoritarian (i) more receptive to a treatment rationale, (ii) accept the authority of the therapist (who was by then discovered to be only a masters student), (iii) follow the therapist's suggestions and assignment, and (iv) expect to benefit from treatment as the authority figure (the therapist) had stated; to a greater extent than subjects who scored low on the E-NZ scale. However, in the long term this type of halo effect would disappear as the experiences of the more authoritarian subjects would disconfirm their beliefs in their elevated degree of freedom from test anxiety. In other words, the true merit of a treatment would be more clearly appreciated by the subjects as time went on.

Studying habits had a substantial facilitating effect on treatments' effects at decreasing general anxiety; however the more influential variable in this aspect of the study was ITA or state test anxiety which acted to inhibit such course of treatment.

The strong inhibiting effect of general anxiety on the treatment effectiveness in decreasing test anxiety was

greatest by far on the ITA (c.f. Table 6.12.2.3). It seems that pre-treatment levels of state test anxiety and general trait anxiety determine to a large extent each other's post-treatment levels.

7.14.2 Test anxiety self-efficacy

The self-efficacy aspect of this investigation turned out very much as predicted.

A conceptual distinction has been made in this area of research; two measures that would allow the testing of self-efficacy type hypotheses within the conceptual distinctions have been constructed and validated. Moreover, they appear to be better measures than existing and well validated test anxiety measures. Finally, treatment effectiveness paralleled subjects' enhanced SEMA feelings. In short, the distinction here made and the results obtained tie in well with past self-efficacy research (c.f. Bandura, 1982).

7.14.3 Effective Evaluation: Key to Progress

Although the main hypothesis of this study was born out of the data here collected, the importance of devising a sensitive evaluative framework cannot be overemphasized. Needless to say, the superiority of the PK treatment was identified thanks to the effectiveness of the evaluative framework in doing so. We have seen that when the traditional evaluative procedures were employed the superiority of the PK

group was masked.

Of course, the merits of the framework for evaluation here developed need to be replicated. Some aspects have to be further refined; for instance, more treatment moderating factors are probably waiting to be discovered. Some other aspects also need to be further validated. For example, the stronger correlation between test anxiety and academic performance in the broad middle range of numerical ability needs to replicated, and so does the positive correlation between test anxiety changes and changes in scores on academic ability tests together with the negative correlation between test anxiety changes and academic performance changes (the author has already made plans to this end). In sum, the two goals achieved by this study are both very important and highly interrelated.

The reader might have realized by now that the results from this study have strong implications relating to the rest of the test anxiety treatment literature. Very early on in the short history of this research, attention has been directed towards establishing which form of a treatment is more effective This task turned out to be a very elusive one, as has been argued in Chapter II, inconsistent results making it impossible to draw confident conclusions.

The reason for this problem, as has been discussed in section 2.3, lies in the shortcomings of previous evaluative procedures. In the light of the results here obtained there are questions that await an answer. When a significant between-group difference on test anxiety decrements was found, was this due to the differential effectiveness of the treatments or was it an artifact resulting from the different

personality characteristics of subjects in the various groups? When a treatment group produced significant improvements in academic performance, was it due to treatment or was it due to the fact that, unlike the other treatment groups, it was composed of subjects with a numerical ability of average levels? The converse of the above questions also await answers.

The results from this study could provide the tools for resolving the inconsistencies of previous research. No doubt the evaluation framework here devised is complex but test anxiety is a complex phenomenon and the evaluation of a test anxiety treatment ought to be just as complex.

At this point I call for researchers in this topic to make better use of their data. A test anxiety treatment study is a great opportunity for studying anxiety when it is in a dynamic state. Not only is it interesting to study the influences of personality factors upon the dynamics of test anxiety, it is also important that we investigate such influences if we want to compare different treatments.

CONCLUSIONS

When I first approached Roger Katz and told him that I was interested in doing a thesis on the management of problems students encounter at university, he replied that I ought to narrow my topic, and mentioned test anxiety. It was the first time these two words reached my ears contiguously. Initially this thesis was going to be far simpler and more straightforward, but the complexity of the topic interacted with the characteristics of the author in a rather deterministic fashion to dictate to a large extent the end product. The virtually unlimited freedom allowed by the supervisory style of my supervisors has greatly facilitated the actualization of this thesis in the form that it is.

In many ways this project can be thought of as an exercise towards "giving psychology away", an endeavour I had been committed to long before I came across George Miller's (1969) presidential address to the American Psychological Association. One of the achievements of the present study lies in showing that paraprofessionals can do much towards the challenge of giving psychology away.

Although I may only speak for myself and not for the professionals, particularly during the 4 weeks when treatment sessions took place I learned a great deal about the practice of psychology which in turn has helped me understand psychology and people a great deal better. In awareness I hoped for and achieved drastic changes in the way this study's subjects reacted to examinations and other evaluative situations, but I don't believe they were aware of the changes

that were happening in my appreciation of the practice of psychology, and above all the appreciation of whole people. So the challenge of giving psychology away to the people has a dual purpose for paraprofessionals: they help and they are being helped ... they help each other!

The target treatment has a skills improvement component and an anxiety management component and from this study and past research it appears that such a two component treatment is more effective than either single component treatments. It is hoped that this thesis will help to bridge the gap between the skills training approach (e.g., Larson, 1984) and the anxiety management approach (e.g., Meichenbaum, 1977). A step further along this dimension, is it too much to hope that this thesis may help decrease ridges in the applied fields of psychology which are caused by different and contending schools? I do believe that in time the practice of psychology will resolve these types of longstanding antagonisms: we cannot afford to turn a blind eye to techniques that achieve results. But these concerns may be beyond the scope of this thesis.

While test anxiety treatment research has been used as a convenient arena for testing psychotherapy techniques, the force of this study works in the opposite direction: better understanding of the problem of test anxiety and its treatment may give insight into better ways to solve other similar problems. For instance, social anxiety and unassertiveness, which I believe are strongly linked, could be better tackled by the double barrel approach which this study has tested, namely one shot at the anxiety aspect (i.e. by employing an anxiety reduction technique) and the other at the skills

aspect (i.e., by adding social skills training or assertiveness training skills). A corollary to these propositions is the belief that test anxiety and other forms of specific anxiety differ only in the person-situation interaction that produces anxiety, not the nature of the specific anxieties. I believe not only that test anxiety is very complex but also that other specific anxieties are very complex phenomena. Because test anxiety is such a heavily researched topic the knowledge gained in this area can be "translated" to apply to other forms of specific anxiety which have been less amenable to the social scientist.

Nonetheless, I believe that test anxiety is the most complex and the most fascinating of all specific anxieties: what makes it so is the fact that academic achievement is such a salient and integral part of modern western societies.

This project is tantamount to a vigorous effort towards the understanding of test anxiety treatment as Meichenbaum and Butler (1980) have called for. It has reviewed the past test anxiety treatment literature more thoroughly and more methodically than past reviews, and it has devised a treatment and an evaluative framework strongly anchored to past research, thereby allowing future evidence on the topic to be added in a cumulative fashion. In a sense, it has helped organize and build upon a body of research woven into a tapestry of picturesque disorder.

In the end its long term achievement will depend on the recognition of the importance of the evaluative framework together with the recognition of the powerful rationale behind the target treatment here developed.

REFERENCES

Adorno, T. W., Frenkel-Brunswick, E., Levinson, D. J., & Sanford, R. N. The Authoritarian Personality. W. W. Norton and Company, New YorK, 1969.

Allen, G. J. Effectiveness of study counseling and desensitization in allleviating test anxiety in college students. Journal of Abnormal Psychology, 1971, 77, 282-289.

Allen, G. J. The Behavioural Treatments of Test Anxiety. Recent research and future trends, Behaviour Therapy, 1972, 3, 253-262.

Allen, G. J. Treatment of test anxiety by group-administered and self-administered relaxation and study counselling. Behaviour Therapy, 1973, 4, 349-360.

Allen, G. J. & Desaulniers G. J. Effectiveness of Study Counselling and desensitization in alleviating test anxiety: A two-year follow-up. Journal of Abnormal Psychology, 1974, 83, 186-191.

Allen, G. J., Elias, M. J. & Zlotlow, S. F. Behavioural interventions for alleviating test anxiety: A methodological overview of current therapeutic practices. In I. G. Sarason (Ed.), Test anxiety: Theory, research and applications. Hillsdale, N. J. : Erlbaum, 1980.

Allen, G., Lerner, W. M., & Hinrichsen, J. J. Study Behaviours and their relationships to test anxiety and academic performance. Pyschological Reports, 1972, 30, 407-410.

Alpert, R. , & Haber, R. N. Anxiety in academic acheivement

situations. Journal of Abnormal and Social Psychology, 1960, 61 207-215.

Altmaier, E. M., & Woodward, M. Group Vicarious Desensitization of Test Anxiety. Journal of Counselling Psychology, 1981, 28, 467-469.

Anastasi, A. Pyschological testing. 3d ed. New York: McMillan, 1968.

Applezwig, D. G. Some determinants of behavioural rigidity Journal of Abnormal and Social Psychology, 1954, 49, 224-228.

Bandura, A. Self-efficacy mechanism in human agency American Pyschologist, 1982, 37, 122-147.

Bandura, A. Self-efficacy: Toward a unifying theory of behavioural change. Psychological Review, 1977 84, 191-215.

Barabasz, A. F. & Barabasz, M. Effects of Rational-emotive therapy on psychophysiological and reported measures of test anxiety. Journal of Clinical Psychology, 1981, 37, 511-514.

Bass, B. M. Authoritarianism or acquiescence? Journal of Abnormal and Social Psychology, 1955, 51, 616-623.

Beneke, W. N., & Harris M. B. Teaching self-control of study behaviour. Behaviour Research and Therapy, 1972, 10, 35-42.

Bistline, J. L., Jaremko M. E., & Sobleman S. The relative contributions of covert reinforcement and cognitive restructuring to test anxiety

reduction. Journal of Clinical Psychology, 1980, 36, 723-728.

Bloch, S. & Brackenridge, C. J. Pyschological performance and biochemical factors in medical students under examination on stress. Journal of Pyschosomatic Research, 1972, 16, 25-33.

Bloom, M. H., & Segal, J. F. The use and effectiveness of systematic desensitization and study skills method in a test anxiety reduction program. Paper presented at the annual meeting of the Western College Reading Association, 1977.

Bruch, M. A. Relationship of test-taking strategies to test anxiety and performance: Towards a task analysis of examination behaviour. Cognitive Therapy and Research, 1981, 5, 41-56.

Brown, C. H. Emotional reactions before examinations: II. Results of a questionnaire. Journal of Psychology, 1938(a), 5, 11-26.

Brown, C. H. Emotional reactions before examinations: III. Intercorrelations. Journal of Psychology, 1938(b), 5, 27-31.

Brown, C. H. & Gelder, D. V. Emotional reactions before examinations: I. Physiological changes. Journal of Psychology, 1938, 5, 1-9.

Brown, R. W. Social Psychology. New York: Free Press, 1965.

Brown, R. W. A determinant of the relationship between rigidity and authoritarianism. Journal of Abnormal and Social Psychology, 1953, 48, 469-476.

Brown, W. H., & Holtzman, W. H. Survey of study habits and attitudes. New York: The Psychological Corporation, 1966.

Boor, M. Relationship of test anxiety and academic performance when controlled for intelligence. Journal of Clinical Psychology, 1972, 28, 171-172.

Boor, M. Test Anxiety and classroom examination Performance: A reply to Daniels and Hewitt (1978). Journal of Clinical Psychology, 1980, 36, 177-179.

Boutin, G. E., & Tosi, D. J. Modification of irrational ideas and test anxiety through rational stage directed hypnotherapy [RSOH]. Journal of Clinical Psychology, 1983, 39, 382-391.

Brehm, J. W. A theory of psychological reactance. New York: Academic Press, 1966.

Carrier, N. A. and Jewell, D. O. Efficiency in measuring the effects of anxiety upon academic performance. Journal of Educational Psychology, 1966, 57, 23-26.

Carron, A. J. Curiosity, acheivement, and avoidant motivation as determinants of epstemic behaviour. Journal of Abnormal and Social Psychology, 1963, 67, 535-549.

Cattell, R. B., & Scheier, I. H. The nature of anxiety: a review of thirteen multivariate analyses comprising 814 variables. Psychological Reports, 1958, 4, 351-388.

Cook, D. C. & Campbell, D. T. Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally, 1979.

Cooley, E. J., & Spiegler, M. D. Cognitive versus emotional coping responses as alternative to test anxiety. Cognitive Therapy and Research, 1980, 4, 159-166.

Cooley, W. W., & Lohnes, P. R. Multivariate procedures for the behavioural sciences. New York: Wiley, 1962.

Cooperative School and College Ability Tests - Series II Educational Testing Service, 1967.

Cowell, M. D., & Entwistle, N.J. Relationship between personality, study attitudes and academic performance in a technical college. British Journal of Educational Psychology, 1971, 41, 85-89.

Cormier, W. H., & Cormier, L. S. Interviewing strategies for helpers : a guide to assessment treatment and evaluation. Montenerey, Calif.: Brooks/Cole, 1979.

Cornish, R. D., & Dilley, J. S. Companion of three methods of reducing test anxiety: Systematic desensitization, implosive therapy, and study counselling. Journal of Counselling Psychology, 1973, 20, 499-503.

Cronbach, L. J., & Furby, L. How should we measure "change" or should we? Psychological Bulletin, 1971, 74, 68-80.

Crossley, T. The examination and validation of attentionally based test anxiety reduction on college

freshemen. Unpublished M.A. thesis, University of New Brunswick, 1977 (cited in Wine, 1980).

Culler, R. E. & Holahan, C. J. Test Anxiety Performance: The effects of study-related Behaviours. Journal of Educational Psychology, 1980, 72, 16-20.

D'Alelio, W. A., & Murray, E. J. Cognitive Therapy of Test Anxiety. Cognitive Therapy and Research, 1981, 5, 299-307.

Daniels, B., & Hewitt, J. Anxiety and classroom examination performance. Journal of Clinical Psychology, 1978, 34, 340-345.

Deese, J., & Deese, E. K. How to study. New York: McGraw Hill, 1979.

Deffenbacher, J. L. Relationship of worry and emotionality to performance on the Miller Analogies Test. Journal of Educational Psychology, 1977, 69, 191-195.

Deffenbacher, J. L. Worry, emotionality and task generated interference in test anxiety: Analogies Test. Journal of Educational Psychology, 1978, 69, 191-195.

Deffenbacher, J. L. Worry and Emotionality in test anxiety. In I. G. Sarason (Ed.), Test anxiety: Theory, research, and applications. Hillsdale, N. J.; Erlbaum, 1980.

Deffenbacher J. L., Deitz S. R., & Hazaleus S. L. Effects of Humour and Test anxiety on Performance, Worry and Emotionality in naturally occurring exams. Cognitive Therapy and Research, 1981, 5, 225-228.

Deffenbacher, J. L., & Hahnloser, R. M. A comparison of cognitive restructuring and progressive relaxation in test anxiety reduction. Unpublished manuscript, Colorado State University, 1978. (cited in Deffenbacher, 1980).

Deffenbacher, J. L. & Hahnloser R. M. Cognitive and Relaxation coping skills in stress inoculation. Cognitive Therapy and Research, 1981, 5, 211-215.

Deffenbacher, J. L. & Hazaleus, S. L. Cognitive enotional, and physiological components of test anxiety. Cognitive Therapy and Research, 1985, 9, 169-180.

Deffenbacher, J. L., Michaels, A. C., Michaels, T., & Daley, P. C. Comparison of Anxiety Management Training and Self-Control Desensitization. Journal of Counselling Psychology, 1980, 27, 232-239.

Deffenbacher, J. L., Mathis, H., Michaels, A. C. Two Self-Control Procedures in the Reduction of Targeted and Non-targeted Anxieties. Journal of Counselling Psychology, 1979, 6, 120-127.

Deffenbacher, J. L., & Parks, D. H. A comparison of Traditional and Self-Control Desensitization. Journal of Counselling Psychology, 1979, 26, 93-97.

Delmonte, M. M. & Ryan G. M. The Cognitive-Somatice Anxiety Questionnaire (CSAQ): A factor analysis. British Journal of Clinical Psychology, 1983, 22, 209-212.

Dember, W. N., Nairne, F., & Miller, F.J. Further validation of the Alpert-Haber Acheivement Anxiety Test. Journal of Abnormal and Social Psychology, 1962, 65, 427-428.

Denney, D. R. Self-control approaches to the treatment of test anxiety. In I. G. Sarason (Ed.), Test anxiety: Theory, research, and applications. Hillsdale, N. J.: Erlbaum, 1980.

Denney D. R. & Rupert, P. A. Desensitization and self-control in the treatment of test anxiety. Journal of Counselling Psychology, 1977, 24, 272-280.

Denney J. P. Effects of anxiety and intelligence on concept formation. Journal of Experimental Psychology, 1966, 72, 596-602.

Desiderato, O., & Koskinen, P. Anxiety, study habits and academic acheivement. Journal of Counselling Psychology, 1969, 16, 162-165.

Dickinson, G. L. Outcome efficacy of a cognitive-behavioural treatment for test anxiety as a function of level of self-control. Diss Abst. Int. 1983, May Vol 43(11-13) 3729 ISSN: 041942009.

Dillard, John; Warrior, Benjamin; Jaquilen-Perrin, D. W. Efficacy of test-wiseness on test anxiety and reading acheivement among black youth. Psychological Reports, 1977, 41, 1135-1140.

Di Tomasso, R. A. A meta-analysis of therapy outcome research with test anxieties college students (Doctoral dissertation University of Pennsylvania, 1980). Dissertation

Abstracts International, 1981, 41, 3884B.

Doctor, R. M. & Altman, F. Worry and emotionality as components of test anxiety: Replication and further data. Psychological Reports, 1969, 24, 563-568.

Doctor, R. M., Apomte, J., Burry, A. & Welch, R. Group counselling versus behaviour therapy in treatment of College underacheivement. Behaviour Research and Therapy, 1970, 8, 87-90.

Doris, J., & Sarason, S. Test anxiety and blame assignment in a failure situation. Journal of Abnormal and Social Psychology, 1955, 50, 335-338.

Ellis, A. Reasons and emotions in Psychotherapy. New York: Lyle Stuart, 1962.

Entwistle, D. R. Evaluations of study-skills courses: A review. Journal of Educational Research, 1960, 53, 243-251.

Entwistle, N. J. & Entwistle D. The relationship between personality, study methods and academic performance. British Journal of Educational Psychology, 1970, 40, 132-142.

Entwistle, N. J., & Wilson J. D. Personality, study methods and academic performance. University Quarterly 1970, 24, 147-156.

Fabick S. D. The relative effectiveness of systematic desensitization, cognitive modification, and mantra meditation on the reduction of test anxiety (Doctoral dissertation, West Virginia University, 1976). Dissertation Abstracts International, 1977, 37, 4862A. University microfilms No DBJ77-02548.

Finger, R. & Galassi J. P. Effects of molifying cognitive versus emotionality responses in the treatment of test anxiety. Journal of Consulting and Clinical Psychology, 1977, 45, 280-287.

Fishman, J. A., & Panasella, A. K. College admission - selection studies. Review of Educational Research, 1960, 30, 298-310.

Fox, L. Effecting the use of efficient study habits. In R. Ulrich, T. Stachnik, and J. Mabry (Eds.), The control of human behaviour (Vol.1) Scott, Forsman, 1966.

Fox, L. H. Cooperative School and College Ability Tests, Series II. In O.K. Buros (Ed.). The eighth mental measurements yearbook, 1978. Highland Park, N. J.: The Gryphon Press, 1978.

Frenkel-Brunswick, E. Intolerance of ambiguity as an emotional and perceptual personality variable. Journal of Personality, 1949, 18, 108-143.

Frost, B. P. Studies of Anxiety and Educational Acheivement. Revista Interamericana de Psicologia, 1969, 3, 83-91.

Galassi, J. P., Frierson, H. T. Jnr., Siegel R. G. Cognitions test anxiety and test performance: A closer look. Journal of Consulting and Clinical Psychology, 1984, 52, 319-320.

Galassi, J. P.; Frierson, H. T. Jnr.; & Sharer, R. Behaviour of High, Moderate, and Low Test Anxious students during an actual test situation. Journal of consulting and Clinical Psychology,

1981(a), 49, 51-62.

Galassi, J. P., Frierson, H. T. Jnr.; Sharer, R. Concurrent versus Retrospective Assessment in Test Anxiety Research. Journal of Counselling and Clinical Psychology, 1981(b), 49, 614-615.

Ganzer, V. J. Effects of audience presence and test anxiety on learning and retention in a serial learning situation. Journal of Personality and Social Psychology, 1968, 8, 194-199.

Geen, R. G. Test anxiety observation, and range of cue utilization. British Journal of Social and Clinical Psychology, 1976, 15, 253-259.

Geen, R. G. The effects of anticipation of positive and negative outcomes on audience anxiety. Journal of Consulting and Clinical Psychology, 1977, 45, 715-716.

Gaudry, E. & Bradshaw G. D. The differential effect of anxiety on performance in progressive and terminal school examinations. Australian Journal of Psychology, 1970, 22, 1-4.

Gaudry, E., & Fitzgerald, D. Test anxiety, intelligence, and academic acheivement. In Gaudry, E. and Spielberger, C. D. Anxiety and educational achievement. Wiley, Sydney, 1971.

Gaudry, E., & Spielberger, C. D. Anxiety and intelligence in paired-associate learning. Journal of Educational Psychology, 1970, 61, 386-391.

Goldiamond I. Self-control procedures in personal behaviour problems. In R. Ulrich, T. Stachnik, and J. Mabry (Eds.), The control of human behaviour (Vol. 1).

Scott, Foresman, 1966.

Goldfried, M. R., Decentesco, E. T., Weinberg, L. Systematic rational restructuring as a self-control technique. Behaviour Therapy, 1974, 5, 247-254.

Goldfried, M. R., Linehan, M. M., & Smith, J. L. Reduction of test anxiety through cognitive restructuring. Journal of consulting and Clinical Psychology, 1978, 46, 32-39.

Goldfried, M. R., & Sobocinski, D. The effect of irrational beliefs on emotional arousal. Journal of Consulting and Clinical Psychology, 1975, 43, 504-510.

Goodstein, L. D. Intellectual rigidity and social attitudes. Journal of Abnormal and Social Psychology, 1953, 48, 345-353.

Gordon, W. K. Combination of cognitive group therapy and sublimal stimulation in treatment of test-anxious college males. Diss. Abst. Int. 1983, May, Vol. 43 (11-13) 3731 ISSN: 04194209.

Gordon, E. M. & Sarason S. B. The relationship between "test anxiety" and "other anxieties". Journal of Personality, 1955, 23, 317-323.

Gorsuch, R. L. The general factor in the test anxiety questionnaire. Psychological Reports, 1966, 19, 308.

Grooms, R. R., & Endler, N. S. The effect of anxiety on academic achievement. Journal of Educational Psychology, 1960, 51, 299-304.

Hahnloser, R. M. A comparison of cognitive restructuring and progressive relaxation in test anxiety

reduction (Doctoral dissertation, University of Oregon, Eugene, 1974). Dissertation Abstracts International, 1974, 35, 14444A. (University Microfilms No 74-18, 897)

Harper, F. B. W. Specific anxiety theory and the Mandler-Sarason Test Anxiety Questionnaire. Educational and Psychological Measurement, 1971, 31, 1011-1014.

Harris, G. & Johnson, S. B. Comparison of Individualised Covert Modelling, self-control desensitization, and study skills training fo alleviation of test anxiety. Journal of Consulting and Clinical Psychology, 1980, 48, 186-194.

Harris, G. M. & Johnson, S. B. Coping imagery and relaxation instruction in a covert Modelling Treatment for Test Anxiety. Behaviour Therapy, 1983, 14, 144-157.

Heaven, P. C. L. Authoritarianism or acquiesence? South African findings. Journal of Social Psychology, 1983, 119, 11-15.

Heckhausen, H. Task irrelevant cognitions during an exam. Incidence and effects! In H. Krohne and L. Laux (Eds.), Achievement stress and anxiety. Washington, D.C.: Hemisphere, 1982.

Hendel, D. D. Experimental and Affective Correlates of Math Anxiety in Adult Women. Psychology of Women Quarterly, 1980, 5, 219-230.

Himble, D. P., Thyer, B. A., Papsdorf, J. D., Caldwell, S. In vivo distraction coping in the treatment of test anxiety: A 1-year follow-up study.

Journal of Clinical Psychology, 1984, 40, 458-462.

Hodges, W. F. The psychophysiology of Anxiety. In M. Zuckerman and C. D. Spielberger (Eds.), Emotions and anxiety: New concepts, methods, and applications. Lawrence Erlbaum, Hillsdale, New Jersey, 1976.

Hollandsworth, J. G., Glazeski, R. C.; Kirkland, K, Jones, G. E., & Van Norman, L. R. An analysis of the nature and effects of test anxiety: Cognitive, Behavioural, and physiological components. Cognitive Research and Therapy, 1979, 3, 165-180.

Holroyd, K. A. Cognition and Desensitization in the group treatment of test anxiety. Journal of Consulting and Clinical Psychology, 1976, 44, 991-1001.

Holroyd, K. A., & Appel, M. A. Test anxiety and physiological responding. In I. G. Sarason (Ed.), Test anxiety: Theory, research, and applications. Hillsdale, N. J.: Erlbaum, 1980.

Holroyd, K., Westbrook, T., Wolf, M., & Badhorn, E. Performance, cognition and physiological responding in test anxiety. Journal of Abnormal Psychology, 1978, 87, 442-451.

Horne, A. M. Effects of Personality Type in Reducing Specific Anxiety with Behavioural and Psychodynamic Theory. Journal of Counselling Psychology, 1974, 21, 340-341.

Horne, A. M. & Matson, J. L. A comparison of modelling,

desensitization, flooding study skills, and control groups for reducing test anxiety. Behaviour Therapy, 1977, 8, 1-8.

Houston, B. K. Trait Anxiety and Coping Behaviour. In H. W. Krohne and L. Laux (Eds.), Achievement Stress and Anxiety. Hemisphere, Washington D.C., 1982.

Huck, S.W., & Jacko, E. J. Effect of varying the response format of the Alpert-Haber acheivement anxiety test. Journal of Counselling Psychology, 1974, 21, 159-163.

Hussian, R. A. & Lawrence, P. S. The reduction of test, state and trait anxiety by test-specific and generalized stress inoculation training. Cognitive Therapy and Research, 1978, 2, 25-37.

Iga, M. Suicide of Japanese youth. Suicide and Life Threatening Behaviourr, 1979, 11, 17-30.

Jackson, B., & Van Zoost, B. Changing study habits through reinforcement contingencies. Journal of Counselling Psychology, 1972, 19, 192-195.

Jackson, P. F., Reid, N. A., & Croft, A.C. Study Habits Evaluation and Instruction Kit. New Zealand Council for Educational Research, 1979.

James, W. What is an emotion? Mind, 1884, 9, 188-205.

Kane, J. S. Evaluation of organisational training programs. Journal of European Training, 1976, 5, 289-338.

Kaplan, R. M., McCordick S. M., & Twitchell M. Is it the cognitive or the behavioural component which makes Cognitive-behaviour modification effective in test anxiety? Journal of Consulting Psychology,

1979, 26, 371-377.

Katahn M. Interaction of anxiety and ability in complex learning situations. Journal of Personality and Social Psychology, 1966, 3, 475-479.

Katahn, M., & Branham, L. Effects of manifest anxiety on acquisition and generalization of concepts from Hullian Theory. American Journal of Psychology, 1968, 81, 575-580.

Katahn, M., Strenger, S., & Cherry, M. Group counselling and behaviour therapy with test anxious college students. Journal of Consulting Psychology, 1966, 30, 544-549.

Kazdin, A. E., & Wilcoxon, L. A. Systematic desensitization and non-specific treatment effects: A methodological evaluation. Psychological Bulletin, 1976, 83, 729-758.

Kight, H. R., & Sassenrath, J. M. The influence of achievement motivation and test anxiety on performance of programmed learning. Journal of Educational Psychology, 1966 57, 14-77.

Kirkland, K., and Hollandsworth, J. G. Jnr. Test anxiety, study skills, and academic performance. Journal of College Student Personnel, 197920, 431-436.

Kirkland, K., & Hollandsworth, J. G. Jnr. Effective test taking : skills acquisition versus Anxiety-Reduction Techniques. Journal of Consulting and

Clinical Pyschology, 1980, 48, 431-439.

Kirkpatrick, D. L. Evaluation of training. In R. L. Craig and L. R. Bittel (Eds.), Training and Development Handbook. New York: McGraw Hill, 1967.

Lalonde, B. The construction and validation of a measure of self-efficacy. Unpublished Ph.D. dissertation, Ontario Institute for Studies in Education, 1979.

Larson, D. Giving psychology away: The skills training paradigm. In D. Larson (Ed.), Teaching psychological skills: Models for giving psychology away. Montenerey, Calif.: Brooks/Cole, 1984.

Lavigne, J. V. The relative efficacy of cognitive-behaviour rehearsal and systematic desensitization on the treatment of test anxiety. Unpublished doctoral dissertation University of Texas, Austin, Tex (1974), (cited in Denney, 1980).

Leal, L. L., Baxter, E. G., Martin, J., Marx, R. W. Cognitive modification and Systematic Desensitization with test anxious High School Students. Journal of counselling psychology, 1981, 28, 525-528.

Lent, R. W. & Russell R. K. Treatment of test anxiety by cue-controlled desensitization and Study-skills Training. Journal of Counselling Psychology, 1978, 25, 217-224.

Levitt, E. E. The Water Jug Einstellung Test as a measure of rigidity. Psychological Bulletin, 1956, 53, 347-370.

Levitt, E. E. The psychology of anxiety. Staples Press,

London, GB, 1968.

Levitt, E. E. & Zelen, S. L. The validity of the Einstellung test as a measure of rigidity. Journal of Abnormal and Social Psychology, 1953, 48, 573-580.

Libert, R. M. & Morris, L. W. Cognitive and emotional components of test anxiety: A distinction and some initial data. Psychological Reports, 1967, 20, 975-978.

Lin, Y., & McKeachie, W. J. Aptitude, anxiety, study habits and academic achievement. Journal of Counselling Psychology, 1970, 17, 306-309.

Little, S. & Jackson, B. The treatment of test anxiety through attentional and relaxation training. Psychotherapy: Theory, Research and Practice, 1974, 11, 175-178.

Luchins, A. S. Mechanization of Problem solving. Psychological Monographs, 1942, 54, No. 6 (Whole No. 248).

Luchins, A. S. Rigidity and ethnocentrism: a critique. Journal of Personality, 1949, 17, 449-466.

Mandler, G. and Sarason, S. B. A study of anxiety and learning. Journal of Abnormal and Social Psychology, 1952 47, 166-173.

Mandler, G. & Watson, D. L. Anxiety and the interruption of behaviour. In C. D. Spielberger (Ed.), Anxiety and Behaviour. New York, Academic Press, 1966.

Many, M. A., & Many, U. A. The relationship between self-esteem and anxiety in grades four through eight. Educational and Psychological Measurement,

1975, 35, 1017-1021.

Marlet, N.J., & Watson, D. Test anxiety and immediate or delayed feedback in a test-like avoidance task. Journal of Personality and Social Psychology, 1968, 8, 200-203.

Marso, R. N. Test item arrangement, testing time and performance. Journal of Educational Measurement, 1970, 7, 113-118.

McKordick, S. M., Kaplan, R. M., Finn, M. E., Smith, S. H. Cognitive Behaviour Modification and Modelling for test anxiety. Journal of Consulting and Clinical Psychology, 1979, 47, 419-420.

McKordick, S. M., Kaplan, R. M., Smith, S., & Finn, M. E. Variations of cognitive behaviour modification for test anxiety. Psychotherapy: Therapy, Research and Practice, 1981, 18, 170-178.

McManus, M. Group desensitization of test anxiety. Behaviour Research and Therapy, 1971, 9, 51-56.

McMillan, J. R., & Osterhouse, R. A. Specific and generalized anxiety as determinants of outcome with desensitization of test anxiety. Journal of Counselling Psychology, 1972, 19, 518-521.

McReynolds, W. T. and Church, A. Self-control study skills development and counselling approaches to the improvement of study behaviour.
Behaviour Research and Therapy, 1973, 11, 233-235.

Michenbaum, D. H. Cognitive modification of test anxious college students.
Journal of Consulting and Clinical Psychology, 1972, 39, 370-380.

Michenbaum, D. Cognitive Behaviour modification: An integrative approach. New York: Plenum Press, 1977.

Michenbaum, D. & Butler, L. Towards a conceptual model for the treatment of test anxiety: Implications for research and treatment. In I. G. Sarason (Ed.), Test anxiety: Theory, research and applications. Hillsdale, N. J.: Erlbaum, 1980.

Michenbaum, D. H., Gilmore, J. B., & Fedoravicious, A. Grou insight versus Group desensitizastion in treating speech anxiety.
Journal of Consulting and Clinical Psychology, 1971, 36, 410-421.

Melnick, J., & Russell R. W. Hypnosis versus systematic desensitization in the treatment of test anxiety.
Journal of Counselling Psychology, 1976, 23, 291-295.

Milholland, J. E. Note on the further validation of the Alpert-Haber Achievement Anxiety Test.
Journal of Abnormal and Social Psychology, 1964, 69, 236.

Miller, G. Psychology as a means of promoting human welfare.

American Psychologist, 1969, 24, 1063-1075.

Mitchell, K. R., Hall, R. F., & Piatkowska, O. E. A program for the treatment of failing college students. Behaviour Therapy, 1975, 6, 324-336.

Mitchell, K. R. & Ingham, R. J. The effects of general anxiety on group desensitization of test anxiety. Behaviour Research and Therapy, 1970, 8, 69-78.

Mitchell, K. R. & Ng, K. T. Effects of group counselling and behaviour therapy on academic achievement of test anxious students. Journal of Counselling Psychology, 1972, 19, 491-497.

Mitchell, K. R., and Piatkowaska, O. Characteristics associated with underachievement: Targets for treatment. Australian Psychologist, 1974(a), 9, 19-41.

Mitchell, K. R., & Piatkowaska, O. E. The effects of group treatment for college underachievers and bright failing underachievers. Journal of Counselling Psychology, 1974(b), 21, 494-501.

Morris, L. W.; Davis, M. A.; & Hutchinson C. H. Cognitive and emotional components of anxiety: literature review and revised worry-emotionality scale. Journal of Educational Psychology, 1981, 73, 541-555.

Morris, L. W., Finkelstein, C. H., and Fisher, W. R. Components of School anxiety: Developmental trends and sex differences. Journal of Genetic Psychology, 1976, 128, 49-57.

Morris, L. W., & Fulmer, R. S.
Test anxiety (worry and emotionality) changes during academic testing as a function of feedback and test importance.
Journal of Educational Psychology, 1976, 68, 817-824.

Morris, L. W., & Liebert, R. M. The effects of anxiety on timed and untimed intelligence tests: Another look.
Journal of Consulting and Clinical Psychology, 1969, 33, 240-244.

Morris, L. W., & Liebert, R. M. Relationship of cognitive and emotional components of test anxiety to physiological arousal and academic performance.
Journal of Consulting and Clinical Psychology, 1970, 35, 332-337.

Morris, L. W., & Liebert, R. M. Effects of negative feedback, threat of shock, and level of trait anxiety on the arousal of two components anxiety.
Journal of Counselling Psychology, 1973, 20, 321-326.

Morris, L. W., & Perez, T. L. Effects of interruption on emotional expression and performance in a testing situation.
Psychological Reports, 1972, 31, 559-564.

Neale, J. M., & Katahn, M. Anxiety choice and stimulus uncertainty. Journal of Personality, 1968, 32, 235-245.

Nicholls, J. G. When a scale measures more than its name denotes: The case of the Test Anxiety Scale for Children.
Journal of Consulting and Clinical Psychology, 1976,

44, 976-985.

O'Neil, H. F. Jnr., Spielberger, C. D., & Hausen, D. N. Effects of state anxiety and task difficulty on computer assisted learning. Journal of Educational Psychology, 1969, 60, 343-350.

Osarchuk, M. A comparison of a cognitive, a behaviour therapy and a cognitive plus behaviour therapy treatment of test anxious college students. (Doctoral dissertation, Adelphi University, 1974). Dissertation Abstracts International, 1975, 36, 3619B. (University Microfilms No. 76-1425,177).

Osipow, S. H., & Kreinbring, I. Temporal stability of an inventory to measure test anxiety. Journal of Counselling Psychology, 1971, 18, 152-154.

Osterhouse, R. A. Desensitization and study skills training as treatment for two types of test anxious. Journal of Counselling Psychology, 1972, 19, 301-307.

Osterhouse, R. A. Classroom anxiety and examination performance of test anxious students. Journal of Educational Research, 1975, 68, 247-249.

Osterhouse, R. A. Group systematic desensitization of test anxiety. In J. D. Krumboltz and C. E. Thoresen (Eds.), Counselling Methods. New York: Halt, Rinehart and Winston, 1976.

Overall, J. E., & Klett, C. J. Applied multivariate analysis. New York: McGraw-Hill, 1972.

Paul, G. L., & Eriksen, C. W. Efforts of test anxiety on "real life" examinations. Journal of Personality, 1964, 32, 480-494.

Paul, G. L. & Shannon, D. T. Treatment of Anxiety through Systematic desensitization in Therapy groups. Journal of Abnormal Psychology, 1966, 71, 124-135.

Paulman, R. G., & Kennelly, K. J. Test anxiety and ineffective test-taking: Different names same construct? Paper presented at the Annual Meeting of the American Psychological Association (90th, Washington, D.C.), Aug. 23-27, 1982, 28 pages.

Pervin, L. A. Aptitude, anxiety and academic performance: A moderator variable analysis. Psychological Reports, 1967, 20, 215-221.

Ray, J. J. A quick measure of achievement motivation-validated in Australia and reliable in Britain and South Africa. Australian Psychologist, 1979, 14, 337-334.

Ray, W. J., Katahn, M., & Snyder, C. R. Effects of test anxiety on a acquisition, retention and generalization of a complex verbal task in a classroom situation. Journal of Personality and Social Psychology, 1971, 20, 147-154.

Raygor, A. L., & Wark, D. M. System for study. New York: McGraw-Hill, 1970.

Reister, B. W., Stockton, R. A., & Maultsby, M. C. Counselling the Test Anxious An Alternative. Journal of College Student Personnel, 1977,

18, 506-510.

Richards, C. S. Behaviour modification of studying through study skills advice and self-control procedures. Journal of Counselling Psychology, 1975, 22, 431-436.

Richardson, F. C., O'Neil, H. F. Jnr., Whitmore S., Judd, W.A. Factor Analysis of the Test Anxiety Scale and evidence concerning the components of test Anxiety. Journal of Counselling and Clinical Psychology, 1977, 45, 704-705.

Richardson, F. C., & Suinn, R. M. The Mathematics Anxiety Rating Scale: Psychometric data. Journal of Counselling Psychology, 1972, 19, 551-554.

Rimm, D. C., & Masters J. C. Behaviour Therapy: Techniques and empirical findings. New York, Academic Press, 1974.

Robinson, F. P. Effective study. (Rev. ed.) New York: Harper, 1961.

Robinson, F. P. Effective Study (4th ed.) New York: Harper Row, 1970.

Rogers, J. M. Operant conditioning in a quasi therapy setting. Journal of Abnormal and Social Psychology, 1960, 60, 247-252.

Rokeach, M. Generalized mental rigidity as a factor in ethnocenturism. Journal of Abnormal and Social Psychology, 1948, 43, 249-278.

Rokeach, M. Rigidity and ethnocenturism a rejoinder. Journal of Personality, 1949, 17, 467-474.

Rosenbach, J. H. Cooperative school and College Ability Tests, Series II. In O. K. Buros (Ed.), The eighth mental measurements year book. Highland Park, N. J.: The Gryphon Press, 1978.

Russell, R. K. & Lent, R. W. Cue-controlled relaxation and systematic desensitization versus non specific factors in treating test anxiety. Journal of Counselling Psychology, 1982, 29, 100-103.

Sanford N. Authoritarian personality in contemporary perspective. In J. N. Knutson (Ed.), Handbook of political psychology. San Francisco, Calif.: Jossey-Boss, 1973.

Sarason, I. G., & Ganzer, V. J. Anxiety Reinforcement and experimental instructions in a free rehabilitation situation. Journal of Abnormal and Social Psychology, 1962, 65, 300-307.

Sarason, I. G., & Ganzer, V. J. Effects of test anxiety and reinforcement history on verbal behaviour. Journal of Abnormal & Social Psychology, 1963, 67, 87-91.

Sarason, I. G., & Koenig, The relationship of est anxiety and hostility to descriptions of self and parents. Journal of Personality and Social Psychology, 1965, 2, 617-621.

Sarason, I. G., & Palola, E. G. The relationship of test anxiety and general anxiety, difficulty of task and experimental instructions to performance. Journal of Experimental Psychology, 1960, 59,

185-191.

Sarason, I. G., Pederson, A. M., and Nyman, B. Test anxiety and the observation of models. Journal of Personality, 1968, 36, 493-511.

Sarason, I. G., & Stoops, R. Test anxiety and the passage of time. Journal of Consulting and Clinical Psychology, 1978, 1, 102-109.

Sarason, I. The effects of anxiety motivational instructions and failure of serial learning. Journal of Experimental Psychology, 1956, 51, 253-260.

Sarason, I. G. Effect of anxiety on two kinds of motivation instructions on verbal learning. Journal of Abnormal and Social Psychology, 1957(a), 54, 166-171.

Sarason, I. G. Test anxiety, general anxiety, and intellectual performance. Journal of Consulting Pyschology, 1957(b0, 21, 485-490.

Sarason, I. G. Effects of anxiety reassurance, and meaningfulness of material to be learned in verbal learning. Journal of Experimental Psychology, 1958, 56, 472-477.

Sarason, I. G. Intellectual and personality correlates of test anxiety. Journal of Abnormal and Social Psychology, 1959(a), 59, 272-275.

Sarason, I. G. Relationships of measures of anxiety and experimental instructions to word association test

performance.

Journal of Abnormal and Social Psychology, 1959(b), 59, 37-42.

Sarason, I. Empirical findings and theoretical problems in the use of anxiety scales.

Psychological Bulletin, 1960, 57, 403-415.

Sarason, I. G. The effects of anxiety and threat on the solution of a difficult task.

Journal of Abnormal and Social Psychology, 1961(a), 62, 165-168.

Sarason, I. G. Test anxiety and intellectual performance of college students.

Journal of Educational Psychology, 1961(b), 52, 201-206.

Sarason, I. G. Test anxiety and intellectual performance.

Journal of Abnormal and Social psychology, 1963, 66, 73-75.

Sarason, I. G. Verbal learning, modelling, and juvenile delinquency.

American Psychologist, 1968, 4, 254-266.

Sarason, I. G. Experimental approaches to test anxiety: Attention and the uses of information. In C.D. Spielbergber (Ed.), Anxiety current trends in theory and research (Vol. II). New York Academic Press, 1972.

Sarason, I. G. Anxiety and self-preoccupation. In I. G. Sarason and C. D. Spielberger (Eds.), Stress and Anxiety (Vol. 2), Washington D.C. Hemisphere, 1975(a).

Sarason, I. G. Test anxiety and the self-disclosing coping

model. Journal of Consulting and Clinical Psychology, 1975(b), 43, 148-153.

Sarason, I. G. The test anxiety scale: concept and research. In C. D. Spielberger & I. G. Sarason (Eds.), Stress and Anxiety (Vol. 5). Washington D.C.: Hemisphere Publishing Corp., 1978.

Sarason, I. G. Introduction to the study of test anxiety. In I. G. Sarason (Ed.), Test anxiety: Theory, research, and applications. Hillsdale, N. J. : Erlbaum, 1980.

Sarason, S.B. What research ssays about test anxiety in elementary school children. NEA Journal, 1959, 48, 26-27.

Sarason, S. B., Davidson, K. S., Lighthall, F.F., Waite, R. R., & Ruebush, B. K. Anxiety in elementary school children. New York: Wile, 1960.

Sarason, S. B., & Mandler, G. Some correlates of test anxiety. Journal of Abnormal and Social Psychology, 1952, 47, 810-817.

Sarason, S. B., Mandler, G., & Craighill, P. G. The effect of differential instruction on anxiety and learning. Journal of Abnormal and Social Psychology, 1952, 47, 561-565.

Sassenrath, J. M. Test anxiety manifest anxiety and concept learning without awareness. Psychological Reports, 1963, 12, 71-81.

Sassenrath, J. M. A factor analysis of relating scale items on the test anxiety questionnaire.

Journal of Consulting Psychology, 1964, 28, 271-377.

Sassenrath, J. M. Anxiety aptitude, attitude and achievement. Psychology in the Schools, 1967, 4, 341-346.

Sassenrath, J. M., Kight, H. R., & Kaiser, H. E. Relating factors from anxiety scales between two samples. Psychological Reports, 1965, 17, 407-416.

Schwartz, G. E., Davidson, R. J., & Goleman, D. J. Patterning of cogniive and somatic processes in the self-regulation of anxiety: Effects of meditation and exercise. Psychosomatic Medicine, 1978, 40, 321-328.

Scrivner, R. W. Systematic desensitization and cognitive modification with high emotionality and high worry subjects (Doctoral dissertation, University of Texas, Austin, 1974). Dissertation Abstracts International, 1974, 35, 2447B. (University Microfilms No. 74-24, 933).

Sieber, J. E. Defining test anxiety: Problems and approaches. In I. G. Sarason (Ed.), Test anxiety: Theory, research, and applications. Hillsdale, N. J. : Erlbaum, 1980.

Smouse, A. D., & Munz, D. C. Item difficulty sequencing and response style: A follow-up analysis. Educational and Psychological Measurement, 1969, 29, 469-472.

Snyder. A. & Deffenbacher, J. Comparison of relaxation as self-control and systematic desensitization in the treatment of test anxiety. Journal of Consulting

and Clinical Psychology, 1977, 45, 1202-1203.

Spence, J. T., & Spence, K. W. The motivational components of anxiety: Drive and drive stimuli. In C. D. Spielberger (Ed.), Anxiety and Behaviour. New York: Academic Press, 1966.

Spielberger, C. D. The effects of manifest anxiety on the academic achievement of College Students. Mental Hygiene, 1962, 46, 420-426.

Spielberger, C. D. Theory and research in anxiety. In C. D. Spielberger (Ed.), Anxiety and behaviour. New York: Academic Press, 1966(a).

Spielberger, C. D. The effects of Anxiety on complex learning and academic achievement. In C. D. Spielberger (Ed.), Anxiety and behaviour. New York: Academic Press, 1966(b).

Spielberger, C. D. Anxiety as an emotional state. In C. D. Spielberger (Ed.), Anxiety: Current trends in theory and research (Vol.1). New York: Academic Press, 1972(a).

Spielberger, C. D. Conceptual and methodological issues in anxiety research. In C. D. Spielberger (Ed.), Anxiety, Current trends in theory and research (Vol.2). New York: Academic Press, 1972(b).

Spielberger, C. D., Anton, W. D., & Bedell, J. The nature and treatment of test anxiety. In M. Zuckerman & C. D. Spielberger (Eds.), Emotions and anxiety: New Concepts, methods and applications. New York: Wiley, 1975 .

Spielberger C. D., Goodstein, L. D., & Dahlstrom, W. G. Complex insidental learning as a function of anxiety and task difficulty. Journal of Experimental Psychology, 1958, 56, 58-61.

Spielberger, C. D., Gonzalez, H.P., Taylor, C.J., Algaze, B., & Anton W. D. Examination stress and test anxiety. In C. D. Spielberger & I. G. Sarason (Eds.), Stress and Anxiety. Vol.5. Washington, Hemisphere, 1978.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. State-Trait Anxiety Inventory. Consulting Psychologists Press, Palo Alto, 1970.

Spielberger, C. D., and Katzenmeyer, W. G. Manifest anxiety, intelligence and college grades. Journal of Consulting Psychology, 1959, 23, 278.

Spielberger, C. D., & Smith, L. H. Anxiety (drive) stress and serial verbal learning. Journal of Experimental Psychology, 1966, 72, 589-595.

Spielberger, C., & Weitz, H. Improving the academic performance of anxious college freshmen. A group counselling approach to the prevention of underachievement. Psychological Monographs, 1964, 78, (13, Whole No. 590).

Stanford, D., Dember, W., & Standford, F. A children's form of the Alpert-Haber achievement anxiety scale. Child Development, 1963, 34, 1027-1032.

Suinn, R. M. The STABS, a measure of test anxiety for

behaviour therapy: Normative data. Behaviour Reseach and Therapy, 1969, 7, 335-339.

Suinn, R. M., & Hill, H. Influence on the relationship between self-acceptance and acceptancce of others. Journal of Consulting Psychology, 1964, 28, 116-119.

Suinn, R. M., Edie, C. A., Nicoletti, J., and Spinelli, P. r. The MARS, a measure of mathematics anxiety: Psychometric data. Journal of Clinical Psychology, 1972, 28, 373-375.

Taylor, D. W. Acomparison of group desensitization with two control procedures in the treatment of test anxiety. Behaviour Research and Therapy, 1971, 9, 281-284.

Taylor, J. A. A personality scale of manifest anxiety. Journal of Abnormal and Social Psychology, 1953, 48, 285-290.

Thyer, B. A., Papsdorf, J.D., Himle, D. P., McCann, B. S., Caldwell, S., & Wickert, M. In vivo distraction coping in the Treatment of Test Anxiety. Journal of Clinical Psychology, 1981, 37, 754-764.

Tobia, S. Anxiety and instruction. In I.G. Sarason (Ed.), Test anxiety: Theory, research, and applications. Hillsdale, N. J.: Erlbaum, 1980.

Tryon, G. S. The measurement and treatment of test anxiety. Review of Educational Research, 1980, 50, 343-372.

Vagg, P. R. The treatment of test anxiety: A comparison of biofeedback and cognitive coping strategies (Doctoral dissertation, University of Michigan, Ann Harbor, 1977). Dissertation Abstracts Internationa, 1978, 38, 5599B. (University Microfilms No.78-04, 828).

Walsh, R. P. Some correlates of test-taking anxiety. Psychological reports, 1968, 22, 449-450.

Walsh, R. P. Test-taking anxiety and psychological needs. Psychological reports, 1969, 25, 83-86.

Walsh, R. P., Engbretson, R. O., & O'Brien, B. A. Anxiety and test-taking behaviour. Journal of Counselling Psychotherapy, 1968, 15, 572-575.

Walter, D., Denzler, L. S., & Sarason I.G. Anxiety and intellectual performance of high school students. Child Development, 1964, 35, 917-926.

Weiner, M. J., & Samuel, W. The effects of attributing arousal to an external source upon test anxiety and performance. Journal of Social and Psychology, 1975, 96, 255-265.

Wilkins, W. Therapy credibility is not a non-specific event. Cognitive Therapy and Research, 1985, 9, 119-125.

Wine, J. D. Test anxiety abd direction of attention. Psychological Bulletin, 1971, 76, 92-104.

Wine, J. D. Test anxiety and self-attention. Paper presented at the American Psychological Association Annual Conference

(New Orleans, Louisiana, August, 1974).

Wine, J. D. Cognitive-attentional theory of test anxiety. In I. G. Sarason (Ed.), Test anxiety: Theory, research, and applications. Hillsdale, N. J.: Erlbaum, 1980.

Wine, J. D. Evaluation anxiety: A cognitive attentional construct. In H. W. Krohne & L. C. Laux (Eds.), Achievement Stress and anxiety. Washington D.C.: Hemisphere, 1982.

Wise, E. H., & Haynes, S. N. Cognitive treatment of test anxiety: Rational Restructuring versus Attentional training. Cognitive Therapy and Research, 1983, 7, 69-77.

Wittmaier, R. Test anxiety and Study Habits. The Journal of Educational Research, 1972, 65, 352-354.

Wittmaier, B. C. Test anxiety, mood, and performance. Journal of Personality and Social Psychology, 1974, 29, 664-669.

Wolf, M. M. Social Validity: The case for subjective measurement or how applied behaviour analysis is finding its heart. Journal of Applied Behaviour Analysis, 1978, 11, 203-214.

Wolpe, J. Reciprocal inhibition therapy. Stanford, Calif.: Stanford Universiy Press, 1958.

APPENDIX A

The Achievement Anxiety Test (AAT)

NAME:_____

COURSE:_____

FOR EACH STATEMENT PUT A CIRCLE ON ONE OF THE 5 POINTS ON THE CORRESPONDING SCALE AS IT APPLIES TO YOU.

1. Nervousness wile taking an exam or test hinders me from doing well.

Always . . . . . Never

2. I work most effectively under pressure, as when the task is very important

Always . . . . . Never

3. In a course where I have been doing poorly, my fear of a bad grade cuts down my efficiency.

Never . . . . . Always

4. Whem I am poorly prepared for an exam or test, I get upset, and do less well than even my restricted knowledge should allow.

This never happens to me . . . . . This practically always happens to me

5. The more important the examination, the less well I seem to do.

Always . . . . . Never

6. While I may (or may not) be nervous before taking an exam, once I start, I seem to forget to be nervous.

I always forget . . . . . I am always nervous during an exam

APPENDIX A (contd.)

7. During exams or tests, I block on questions to which I know the answers, even though I might remember them as soon as the exam is over.

This always happens to me	I never block on questions to which I know the answers

8. Nervousness while taking a test helps me to do better.

It never helps	It often helps

9. When I start a test, nothing is able to distract me.

This is always true of me	This is not true of me

10. In courses in which the total grade is based main on *one* exam, I seem to do better than other people.

Never	Almost always

11. I find that my mind goes blank at the beginning of an exam, and it takes me a few minutes before I can function.

I almost always blank out at first	I never blank out at first

12. I look forward to exams.

Never	Always

13. I am so tired from worrying about an exam, that I find I almost don't care how well I do by the time I start the test.

I never feel this way	I amost always feel this way

APPENDIX A (contd.)

14. Time pressure on an exam causes me to do worse than the rest of the group under similar conditions.

Time pressure always seems to make me do worse on an exam than others	Time pressure never seems to make me do worse on an exam than others

15. Although "cramming" under pre-examination tension is not effective for most people, I find that if the need arises, I can learn material immediately before an exam, even under considerable pressure, and successfully retain it to use on the exam.

I am always able to use the "crammed" material successfully	I am never able to use the "crammed" material successfully

16. I enjoy taking a difficult exam more than an easy one.

Always	Never

17. I find myself reading exam questions without understanding them, and I must go back over them so that they will make sense.

Never	Almost always

18. The more important the exam or test, the better I seem to do.

This is true of me	This is not true of me

19. When I don't do well on a difficult item at the beginning of an exam, it tends to upset me so that I block on even easy questions later on.

This never happens to me	This almost always happens to me

APPENDIX B

The Inventory of Test Anxiety (ITA)

NAME:_____

Read each of the following statements carefully. In the space before each item, indicate how you *actually* felt during your last examination. Use the following scale:

1. The statement did *not* describe my feeling or condition.
2. The feeling or condition was barely noticeable.
3. The feeling or condition was moderately intense.
4. The feeling or condition was strong.
5. The feeling or condition was very strong.

_____ 1. I felt panicky while taking this examination.

_____ 2. I felt during this examination that I wouldn't be able to finish the examination on time.

_____ 3. My mouth got dry during this examination.

_____ 4. Prior to taking this examination, I felt that other students were better prepared for this examination than I was.

_____ 5. My mind went blank at the beginning of this examination. It took me a few minutes to function.

_____ 6. I feel that I let myself and other persons down by my performance on this examination.

_____ 7. I felt my heart beating fast during this examination.

_____ 8. I found myself worrying about a low grade before this examination.

_____ 9. During this examination, I found myself thinking about the consequences of failure.

_____ 10. I got so tense during this examination that my stomach became upset.

_____ 11. After finishing this examination, I feel that I could have done better than I actually did.

_____ 12. I got a headache during this examination.

_____ 13. While taking this examination, I found myself thinking of how much brighter other students are than I am.

_____ 14. My hands perspired during this examination.

_____ 15. I did not feel very confident of my performance before I took this examination.

_____ 16. I got so nervous during this examination that I forgot facts which I really knew.

APPENDIX C

The Test Anxiety Self-Efficacy Questionnaire (TASEQ)

The attached form lists activities that could be performed in the different problem areas. Under the column Can Do, check (✓) the tasks you expect you could do if you were asked to perform them now.

For the tasks you check under Can Do, indicate in the column Confidence how confident you are that you could do them. Rate your degree of confidence by recording a number from 10 to 100 using the scale given below:

10	20	30	40	50	60	70	80	90	100
quite uncertain				moderately certain					certain

Remember, rate what you expect you could do and your confidence if you were asked to perform the tasks now.

APPENDIX C (contd.)

Practice Rating

To familiarize you with the rating form, please complete this practice item first.

Confidence Scale

10	20	30	40	50	60	70	80	90	100
quite uncertain				moderately certain					certain

PHYSICAL STRENGTH	Can do	Confidence
Lift a 10 pound weight	_____	_____
Lift a 15 pound weight	_____	_____
Lift a 20 pound weight	_____	_____
Lift a 30 pound weight	_____	_____
Lift a 40 pound weight	_____	_____
Lift a 50 pound weight	_____	_____
Lift a 60 pound weight	_____	_____
Lift a 70 pound weight	_____	_____
Lift an 80 pound weight	_____	_____
Lift a 100 pound weight	_____	_____
Lift a 120 pound weight	_____	_____
Lift a 140 pound weight	_____	_____
Lift a 160 pound weight	_____	_____
Lift a 180 pound weight	_____	_____

APPENDIX C (contd.)

Suppose you were about to sit a *scheduled* exam or test. Check (✓) those items you could do and rate your confidence level.

Confidence Scale

10 20 30 40 50 60 70 80 90 100
quite uncertain — moderetely certain — certain

During the exam/test:	Can do	Confidence
I would feel I will do alright on this exam	_____	_____
I would feel in control of my reactions	_____	_____
I would feel the grade I'll gain will reflect my study efforts	_____	_____
I would feel that my course grade will stay the same or increase after this exam	_____	_____
I'll be able to concentrate as I take this exam	_____	_____
My mind will be clear as I take this exam	_____	_____
If I'll have trouble with one question I will feel that this will not affect my performance on other questions	_____	_____
I will feel that time pressure affects me the same as it affects others	_____	_____
I would feel as bright or brighter than other students	_____	_____
If I will begin to feel the exam is hard I would do my best without worrying about it	_____	_____
I would never think how awful it would be if I failed or did poorly on the exam	_____	_____
I will remember all the relevant facts I know	_____	_____
I would never call myself "stupid"	_____	_____
The night before the exam I would have slept as usual	_____	_____
I would do well in spite of feeling nervous during the exam	_____	_____
I would do well in spite of my jitters	_____	_____
I would do well in spite of feeling panicky	_____	_____
My mind will never go blank	_____	_____

Confidence Scale

10 20 30 40 50 60 70 80 90 100
quite uncertain — moderately certain — certain

During the exam/test:	Can do	Confidence
I would do well in spite of my sweating palms	_____	_____
I would do well in spite of my heart beating fast	_____	_____
I would never worry too much once the exam is over	_____	_____

APPENDIX D

The Post Test Anxiety Self-Efficacy Questionnaire (P-TASEQ)

Complete this forms as soon as you can after your first/[redacted] exam.

Date:_____ Time:_____ Name:_____

This form has a list of feelings, thoughts, and reactions relating to taking exams. Under "Did" check (✓) those items that describe your feelings, thoughts, and reactions during the exam you have just taken. For those items you check under "Did" rate the degree to which the statement accurately describes your experience during the exam under the "Accurate" column.

Rate the degree of accuracy of each item by using the scale below:

Accuracy Scale

10 20 30 40 50 60 70 80 90 100

quite inaccurate — moderately accurate — accurate

	Did	Accuracy
I felt I was going to do alright on this exam	_____	_____
I felt in control of my reactions	_____	_____
I felt that the grade I'll gain will reflect my study efforts.	_____	_____
I felt that my course grade will stay the same or increase after this exam.	_____	_____
I was able to concentrate as I took this exam	_____	_____
I had trouble with one question but I felt that was not going to affect my performance on other questions	_____	_____
I felt that time pressure was affecting me the same as it was affecting the others	_____	_____
I felt as bright or brighter than other students	_____	_____
I began to feel the exam was hard but I went on to do my best without worrying about it	_____	_____
I never thought how awful it would be if I failed or did poorly on the exam	_____	_____
I remembered all the relevant facts I know	_____	_____
I never called myself "stupid"	_____	_____
The night before the exam I slept as usual	_____	_____
I did well inspite of feeling nervous during the exam	_____	_____

APPENDIX D (contd.)

Accuracy Scale

10	20	30	40	50	60	70	80	90	100
quite inaccurate				moderately accurate					accurate

	Did	Accuracy
I did well inspite of my jitters	_____	_____
I did well inspite of feeling panicky	_____	_____
My mind never went blank	_____	_____
I did well inspite of my sweating palms	_____	_____
I did well inspite of my heart beating fast	_____	_____
I am not going to worry too much about this exam now	_____	_____

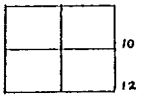
APPENDIX E

The Cognitive-Somatic Anxiety Questionnaire (CSAQ)

REACTIONS TO ANXIETY

Rate the degree to which you typically experience each of the following reactions when you are feeling ANXIOUS by circling the appropriate number.

		Not at all	A little	Somewhat	Quite a lot	Very much
1.	I become immobilized	1	2	3	4	5
2.	I worry too much over something that does not really matter	1	2	3	4	5
3.	My heart beats faster	1	2	3	4	5
4.	I imagine terrifying scenes	1	2	3	4	5
5.	I get diarrhea	1	2	3	4	5
6.	I find it difficult to concentrate because of uncontrollable thoughts	1	2	3	4	5
7.	I pace nervously	1	2	3	4	5
8.	Some unimportant thought runs through my mind and bothers me	1	2	3	4	5
9.	I can's keep anxiety provoking pictures out of my mind	1	2	3	4	5
10.	I feel tense in my stomach	1	2	3	4	5
11.	I can't keep anxiety provoking thoughts out of my mind	1	2	3	4	5
12.	I perspire	1	2	3	4	5
13.	I feel like I am losing out on things because I can't make up my mind soon enough	1	2	3	4	5
14.	I feel jittery in my body	1	2	3	4	5

C  10

S 12

Schwatz et al, 1978.

APPENDIX F

The Ray Achievement Motivation (RAM) scale

Name : _____

		Yes	Don't Know	No
1.	Is being comfortable more important to you than getting ahead?			
2.	Are you satisfied to be no better than most other people at your job?			
3.	Do you like to make improvement to the way the organisation you belong to functions?			
4.	Do you take trouble to cultivate people who may be useful to you in your career?			
5.	Do you get restless and annoyed when you feel you are wasting time?			
6.	Have you always worked hard in order to be among the best in your own line? (school, organisation, profession).			
7.	Would you prefer to work with a congenial but incompetent partner rather than with a difficult but highly competent one?			
8.	Do you tend to plan ahead for your job or career?			
9.	Is "getting on in life" important to you?			
10.	Are you an ambitious person?			
11.	Are you inclined to read of the successes of others rather than do the work of making yourself a success?			
12.	Would you describe yourself as being lazy?			
13.	Will days often go by without your having done a thing?			
14.	Are you inclined to take life as it comes without much planning?			

APPENDIX G

The Einstellung Water Jug (EWJ) test

Here is a problem solving task for you.

Complete this as quickly as possible.

Time yourself. What is the time now? ____/_____/_____
hour/minutes/seconds

You are to determine how you could measure out various quantities of water using bottles of specific sizes given below.

Each bottle can exactly measure only its full volune as no gradiations are marked.

The best solution is the shortest possible method.

EXAMPLE

Given: Containers of capacities : 31, 61 and 4 pints.
Obtain: 22 pints.
Solution: Fill the bottle that holds 61 pints, from it fill the 31 pint bottle, from the remainder withdraw 4 pints twice.
In short: 61 - 31 - 4 - 4 = 22.

1. Given: Containers of capcities: 39 and 4 pints.
Obtain: 31 pints.
Solution:

2. Given: Containers of capacities: 13, 29, and 3 pints.
Obtain: 10 pints.
Solution:

3. Given: Containers of capacities: 30, 40, and 3 pints.
Obtain: 4 pints.
Solution:

4. Given: Containers of capacities: 31, 61, and 4 pints.
Obtain: 22 pints.
Solution:

5. Given: Containers of capacities: 14, 59, and 10 pints.
Obtain: 25 pints.

APPENDIX G (contd.)

6. Given: Containers of capacities: 23, 49, and 3 pints.
   Obtain: 20 pints.
   Solution:

7. Given: Containers of capacities: 11, 25, and 3 pints.
   Obtain: 8 pints.
   Solution:

8. Given: Containers of capacities: 17, 40, and 6 pints.
   Obtain: 11 pints.
   Solution:

9. Given: Containers of capacities: 10, 23, and 3 pints.
   Obtain: 7 pints.
   Solution:

10. Given: Containers of capacities: 11, 27, and 5 pints.
    Obtain: 6 pints.
    Solution:

11. What is the time now?

    Solution: _____ _____ _____
    hours/minute/seconds

APPENDIX H

The Ethnocentrism-New Zealand (E-NZ) scale

The following statements refer to opinions regarding a number of social groups and issues, about which some people agree and others disagree. Please mark each statement in the left-hand margin according to your agreement or disagreement, as follows:

+1: slight support, agreement −1: slight opposition, disagreement
+2: moderate support, " −2: moderate opposition, "
+3: strong support, " −3: strong opposition, "

____ 1. The many political parties tend to confuse national issues, add to the expense of elections, and raise unnecessary agitation. For this and other reasons, it would be best if all political parties except the two major ones were abolished.

____ Patriotism and loyalty are the first and most important requirements of a good citizen.

____ Certain religious sects whose beliefs do not permit them to salute the flag should be forced to conform to such a patriotic action, or else be abolished.

____ The Maoris would solve many of their social problems by not being so irresponsible, lazy, and ignorant.

____ Any group or social movement which contains many foreigners should be watched with suspicion and, whenever possible, be investigated by the SIS

____ There will always be superior and inferior nations in the world and, in the interests of all concerned, it is best that the superior ones be in control of world affairs.

music and jazz, but it is a mistake to have mixed Negro-white bands.

____ Although women are necessary now in the armed forces and in industry, they should be returned to their proper place in the home as soon as the war ends.

____ Minor forms of military training, obedience, and discipline, such as drill, marching and simple commands, should be made a part of the elementary school educational program.

____ It would be a mistake to have Maoris for foremen and leaders over whites.

____ Present treatment of conscientious objectors, draft-evaders, and enemy aliens is too lenient and mollycoddling. If a person won't fight for his country, he deserves a lot worse than just a prison or a work camp.

____ Maoris may have a part to play in white civilization, but it is best to keep them in their own districts and schools and to prevent too much intermixing with whites.

____ One main difficulty with allowing the entire population to participate fully in government affairs (voting, jobs, etc.) is that such a large percentage is innately deficient and incapable.

____ Manual labor and menial jobs seem to fit the Maori mentality and ability better than more skilled or responsible work.

____ In view of the present unemployment, it is highly important to limit responsible government jobs to native, white, Christian, New Zealanders.

____ In a community of 1,000 whites and 50 Maoris a drunken Maori shoots and kills an officer who is trying to arrest him. The white population should immediately drive all the Maoris out of town.

____ The people who raise all the talk about putting Maoris on the same level as whites and giving them the same privileges are mostly radical agitators trying to stir up conflicts.

____ A large-scale system of sterilization would be one good way of breeding out criminals and other undesirable elements in our society and so raise its general standards and living conditions.

____ The most vicious, irresponsible, and racketeering unions are, in most cases, those having largely foreigners for leaders.

____ There is something inherently primitive and uncivilized in the Maoris as shown in his music and his extreme aggressiveness.

____ We are spending too much money for the pampering of criminals and the insane, and for the education of inherently incapable people.

____ There will always be wars because, for one thing, there will always be races who ruthlessly try to grab more than their share.

____ Most Maoris would become officious, overbearing, and disagreeable if not kept in their place.

APPENDIX I

The First Impression of Treatment (FIT) Questionnaire

Briefly state in your own words what you hope to gain from this program.

Do you think you will?

Would you recommend a friend who is anxious during exams and tests to take part in one such program?

APPENDIX J

The Participant Reactions Questionnaire (PRQ)

Do you think this program has helped you in dealing with your exam anxiety?

Rate the degree to which you think participating in this program has helped you in managing your exam anxiety.

1	2	3	4	5	6	7
Not helped			moderately helped			helped a great deal

Do you think that the therapist has helped you in managing your exam anxiety?

On the scale below, rate the degree to which you think he has helped you manage your exam anxiety.

1	2	3	4	5	6	7
Not helped			moderately helped			helped a great deal

On the scale below, rate the degree to which you think the therapist has understood your particular problem.

1	2	3	4	5	6	7
Not understood			moderately understood			very well understood

On the scale below, rate the therapist expertise judged by you.

1	2	3	4	5	6	7
Not competent			moderately competent			very competent

How authentic do you think this program is? Put a circle on the appropriate point on the scale below.

1	2	3	4	5	6	7
Not authentic			moderately authentic			very authentic

Would you recommend a friend who is text anxious to participate in a program such as this?

Please give your opinion/comments on this program.